Algorithms for Optimizing over Analytic Functions*

Dedicated to Ciprian Foias on occasion of his sixtieth birthday

J. William Helton[†] Orlando Merino[‡] Trent E. Walker[§]

Abstract

Let H_N^∞ be the space of C^N-valued measurable functions on the unit circle ∂D that extend to a uniformly bounded analytic function on the unit disk D. A fundamental optimization problem in frequency domain design where stability is the key constraint are:

$$OPT^\infty \qquad \begin{array}{l} \text{Given a map } \Gamma : \partial D \times C^N \to R^+ \text{ and } E \subset H_N^\infty, \\ \text{minimize } \sup_{e^{i\theta} \in \partial D} \Gamma(e^{i\theta}, f(e^{i\theta})) \text{ over } f \in E \\ \text{and find minimizers } f^* \text{ in } E \text{ if they exist.} \end{array}$$

The main contribution of the article is to present an algorithm for solving OPT^∞ when Γ is smooth and E is the set of functions in H_N^∞ that are smooth on ∂D. The algorithm is quadratically convergent and is conceptually different from existing algorithms. In this article we prove that the algorithm is (locally) quadratically convergent to solutions to OPT^∞ and illustrate practical performance of the algorithm with results of computer runs.

Another problem studied in this article is

$$OPT^1 \qquad \begin{array}{l} \text{Given a map } \Gamma : \partial D \times C^N \to R^+ \text{ and } E \subset H_N^\infty, \\ \text{minimize } \int_0^{2\pi} \Gamma(e^{i\theta}, f(e^{i\theta})) \frac{d\theta}{2\pi} \text{ over } f \in E \\ \text{and find minimizers } f^* \text{ in } E \text{ if they exist.} \end{array}$$

We derive the optimality conditions that characterize local solutions to OPT^1 when $E = A_N$, the set of functions in H_N^∞ that are continuous on ∂D. An algorithm for solving OPT^1 when Γ is smooth and E is the set of smooth H_N^∞ functions is presented and convergence results are derived. Performance of the algorithm is illustrated with results of a computer run.

When $\Gamma(e^{i\theta}, z) = \sum_{j=1}^N p_j(e^{i\theta}) \, |k_j(e^{i\theta}) - z_j|^2$ for (bounded, measurable) complex valued functions k_j and positive valued functions p_j, and $E = H_N^\infty$, the sets

$$S_\omega(c) = \left\{ (z_1, z_2, \ldots, z_N) \in C^N \colon \ \Gamma(e^{i\theta}, z) \le c \right\}$$

are disks in C^N and the problem OPT^∞ has an *explicit solution.* This case and the case when each $S_\omega(c)$ is a matrix ball are very important to engineering. Fortunately, it was solved by Sarason [S], Nagy-Foias [NF], Ando [A], and Adamjan-Arov-Krein [AAK] in papers which originated a vast body of work in both mathematics and engineering.

---

*This work was sponsored by the Air Force Office for Scientific Research and by the National Science Foundation.

[†] Department of Mathematics, University of California, San Diego, California.

[‡] Department of Mathematics, Texas Tech University, Lubbock, Texas.

[§] Department of Mathematics, University of California, Berkeley, California

Contents

1 Introduction

Let H_N^∞ be the space of C^N-valued measurable functions on the unit circle ∂D that extend to a uniformly bounded analytic function on the unit disk D. A fundamental optimization problem in frequency domain design where stability is the key constraint are:

$$OPT^\infty \qquad \begin{array}{l} \text{Given a map } \Gamma : \partial D \times C^N \to R^+ \text{ and } E \subset H_N^\infty, \\ \text{minimize } \sup_{e^{i\theta} \in \partial D} \Gamma(e^{i\theta}, f(e^{i\theta})) \text{ over } f \in E \\ \text{and find minimizers } f^* \text{ in } E \text{ if they exist.} \end{array}$$

The main contribution of the article is to present an algorithm for solving $OPT^\infty$when Γ is smooth and E is the set of functions in H_N^∞ that are smooth on ∂D. The algorithm is quadratically convergent and is conceptually different from existing algorithms. In this article we prove that the algorithm is (locally) quadratically convergent to solutions to $OPT^\infty$and illustrate practical performance of the algorithm with results of computer runs.

Another problem studied in this article is

$$OPT^1 \qquad \begin{array}{l} \text{Given a map } \Gamma : \partial D \times C^N \to R^+ \text{ and } E \subset H_N^\infty, \\ \text{minimize } \int_0^{2\pi} \Gamma(e^{i\theta}, f(e^{i\theta})) \frac{d\theta}{2\pi} \text{ over } f \in E \\ \text{and find minimizers } f^* \text{ in } E \text{ if they exist.} \end{array}$$

We derive the optimality conditions that characterize local solutions to OPT^1 when $E = A_N$, the set of functions in H_N^∞ that are continuous on ∂D. An algorithm for solving OPT^1 when Γ is smooth and E is the set of smooth H_N^∞ functions is presented and convergence results are derived. Performance of the algorithm is illustrated with results of a computer run.

By OPT we will refer to either OPT^∞ or OPT^1.

Motivation.

The OPT problem is central to the design of a system where specifications are given in the frequency domain and stability is a key issue. Suppose our objective is to design a (linear, finite dimensional) system part of which we are forced to use (in control it is called the plant) and part of which is designable.

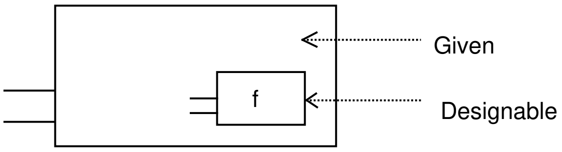

Figure 1: The given part of the system

Suppose that we have available a performance (or cost) function $\Gamma(\omega, f(i\omega))$ of frequency ω and designable part f. Typically Γ is built from the specifications on the design. A measure of goodness of choice of designable f is $||\Gamma(\cdot, f(\cdot))||_p$. One wants to minimize this over all admissible f.

The stipulation that the designable part of the circuit be stable amounts to requiring that f has no poles in the Right Half Plane (R.H.P.). In other words $f \in A_N$ (R.H.P.). This is exactly the R.H.P. version OPT problem. Even when parts of the system other than the designable part are in H^∞ one can frequently reparametrize to get OPT. Consequently OPT arises in a large class of problems.

The special case of OPT^∞ where the sublevel sets of Γ in z,

$$S_\theta(c) = \{z \in C^N : \Gamma(e^{i\theta}, f(e^{i\theta})) \leq c\}$$

are "disks" in C^N is the main mathematical optimization problem in the area called H^∞-control (cf. [Fr]) which is one of the main emphasis of control since the early 1980's.

One reason this subject advanced so quickly is that the mathematics behind these "disk problems" was very well developed. For $N = 1$ this goes back to Pick, to Nevalinna and to Nehari. For $N > 1$ this fell into the province of operator theory and an area called commutant lifting. Fundamental work of Sarason, Nagy-Foias, and Ando gave a powerful theory for solving these problems as did work of Adamjan-Arov-Krein and this type of mathematics formed the corner stone of H^∞-control. Mathematically the richness of the subject is attested to by the fact that numerous books concern it, see Ball-Gohberg-Rodman [BGR], Dym [Dym], Foias-Frazho [F-F], Glover-McFarland [GM], Helton [H], and Young [Yng].

The purely frequency domain algorithms presented in this article apply to a large class of differentiable performance functions Γ. Functions Γ in terms of matrix norms are not differentiable; however, matrix norms can be approximated with smooth norms, possibly allowing frequency domain methods like the ones presented here to be used (this has not been attempted by the authors, nor we know of any efforts in this direction).

Related frequency domain design theories are those originated by Mayne and Polak [MP] and more recently pursued by M. Fan and A. Tits. In a sense the problem they study is more general than OPT; though technically speaking it is not (since it is finite dimensional). They optimize performance over classes of rational functions (which, of course, are highly non-convex). We optimize over the class of all analytic functions on the unit circle and exploit their special structure to obtain the elegant and sharp optimality conditions we use in this paper. Ironically infinite dimensionality of A_N helps because it guarantees that at optimum "every constraint is active". (This is condition I of Theorem 2.2.) In another direction linear programming is a viable approach to OPT problems provided that Γ has sublevel sets $\mathcal{S}_\theta(c)$ which are convex. Then one can approximate the $\mathcal{S}_\theta(c)$ by a simplex and use standard packages. This has been explored extensively by S. Boyd et al. (see [BB] for discussion and references) and crude numerical comparisons of linear programming to other H^∞ methods were done in [BHM], [HS]. The theorems given have required that Γ be a smooth function so they do not apply as stated to a simplex in C^N since a simplex has corners. They would apply to a smooth approximation of the simplex and so they might well be adaptable for use in linear programming based codes. Also for $N = 1$ a general version of Theorem 2.1 appears in [HM] which applies to $S_\theta(c)$ which are simplicies.

Notation.

By L^p_N we represent the set of all C^N-valued measurable functions on the unit circle ∂D having finite norm $||F||_{L^p}$, where as usual,

$$||F||_{L^p} = \left(\int_0^{2\pi} ||F(e^{i\theta})||_{C^N}^p \frac{d\theta}{2\pi} \right)^{1/p} \qquad 1 \le p < \infty$$

and

$$||F||_{L^\infty} = \sup_\theta ||F(e^{i\theta})||_{C^N}$$

By χ we denote the function $\chi(e^{i\theta}) = e^{i\theta}$. Functions $F \in L_N^1$ have a uniquely determined Fourier series expansion (see [Hs])

$$F \sim \sum_{j=-\infty}^{\infty} \hat{F}_j \chi^j$$

where the Fourier coefficients $\hat{F}_j$ are vectors in C^N. By H_N^p we denote the set of all functions $F \in L_N^p$ for which $\hat{F}_j = 0$, for $j = -1, -2, \ldots$. Functions F in H_N^p can be extended to an analytic function on the unit disk D. We will make use of this fact when necessary without more comments, and the function F and its extension to the disk will be refered to under the same name.

We shall also use the notation $H_{N,0}^2 = \chi H_N^2$. If $\mathcal{E}$ is a (closed) subspace of H_N^2, we write $P_{\mathcal{E}}$ for the orthogonal projection operator onto $\mathcal{E}$.

If $E \subset H_N^\infty$, set

$$RE \triangleq \{ f \in E \colon \overline{f(e^{-i\theta})} = f(e^{i\theta}) \quad \text{a.e. on} \quad T \} . \tag{1.1}$$

Another way of defining RE is as the set of elements of E with real Fourier coefficients. Frequency response functions in engineering are all real on the real axis [K]. Thus spaces such as RH_N^∞, RA_N and the corresponding OPT problems are the only ones which occur in engineering applications.

We shall investigate local conditions for solutions to OPT. For this we make use of the terms *optimizer* and of *directional optimizer*, which we now define:

Definition 1 *Let $E \subset H_N^\infty$ and $f^* \in E$. The function f^* is a local optimizer for OPT_E if there exists an open set $V \subset E$ such that $f^* \in V$ and*

$$\sup_\theta \Gamma(e^{i\theta}, f^*(e^{i\theta})) \le \sup_\theta \Gamma(e^{i\theta}, f(e^{i\theta})) \quad (\forall f \in V) . \tag{1.2}$$

The function f^* is a directional optimizer for OPT_E if for each $h \in E$ there exists $t_h > 0$ such that

$$\sup \Gamma(e^{i\theta}, f^*(e^{i\theta})) \le \sup \Gamma(e^{i\theta}, f^*(e^{i\theta}) + t\, h(e^{i\theta})) \quad (\forall t \in (0, t_h)) . \tag{1.3}$$

The terms *strict local optimizer* and *strict local directional optimizer* are defined similarly, with strict inequality in (1.2) and (1.3), respectively. If E is finite dimensional the concepts local optimizer and directional optimizer coincide. This is a simple consequence of the compactness of the unit ball in E. In the infinite dimensional f, a strict directional optimizer may not be a strict local optimizer.

Elements $z \in C^N$ are represented as column vectors. By z^t we denote the transpose of z, so that $\bar{z}^t z = \|z\|^2_{C^N}$. Let G: $C^N \to R$ be a class C^3 function. Then the order 2 expansion of G about z_0 is

$$G(z_0 + w) = G(z_0) + 2\Re \frac{\partial G}{\partial z}(z_0)^t\, w + \bar{w}^t \frac{\partial^2 G}{\partial \bar{z} \partial z}(z_0) w + \Re \left(w^t \frac{\partial G}{\partial z^2}(z_0) w \right) + \mathcal{O}(\|w\|^3_{C^N}) \qquad (1)$$

where $\frac{\partial G}{\partial z}(z_0)$ is a column vector of entries $\frac{\partial G}{\partial z_\ell}(z_0)$, $\ell = 1, ..., N$, and $\frac{\partial^2 G}{\partial \bar{z} \partial z}(z_0)$ (resp. $\frac{\partial G}{\partial z^2}(z_0)$) is an $N \times N$ matrix, with entries $\frac{\partial^2 G}{\partial \bar{z}_\ell \partial z_j}(z_0)$ (resp. $\frac{\partial^2 G}{\partial z_\ell \partial z_j}(z_0)$) $j, \ell = 1, ..., N$. If for each z_0 the matrix $\frac{\partial^2 G}{\partial \bar{z} \partial z}(z_0)$ is strictly positive definite, then G is called *strictly plurisubharmonic*.

2 Characterization of solutions to OPT

2.1 Main results on characterization of solutions to OPT

In this section we present necessary and sufficient conditions for a function f^* to be a solution to OPT^p, for $p = 1$ and $p = \infty$. The result for $p = \infty$ can be found in [HMer2]. The connection between both results is also considered.

Our first theorem gives optimality conditions for OPT^1.

Theorem 1 *Let Γ be of class C^3 in z with all derivatives in z up to order 3 continuous in $e^{i\theta}$. If $f^* \in A_N$ is a local directional solution to OPT^1, then*

I. There exist a function $F \in H^1_N$ such that

$$\frac{\partial \Gamma}{\partial z}(\cdot, f^*) = \chi F$$

II. For all $h \in A_N$,

$$\int_0^{2\pi} \left\{ \overline{h(e^{i\theta})}^t \frac{\partial^2 \Gamma}{\partial z \partial \bar{z}}(e^{i\theta}, f^*(e^{i\theta}))h(e^{i\theta}) + \Re \left(h(e^{i\theta})^t \frac{\partial^2 \Gamma}{\partial z^2}(e^{i\theta}, f^*(e^{i\theta}))h(e^{i\theta}) \right) \right\} \frac{d\theta}{2\pi} \geq 0 \quad (2)$$

Conversely, if f^* satisfies I and II with $>$, then f^* is a strict local directional optimizer for OPT^1.

Proof. For arbitrary $h \in A_N$, $t \in R$ set

$$s_h(t) = \int_0^{2\pi} \Gamma(e^{i\theta}, f^*(e^{i\theta}) + th(e^{i\theta}))\frac{d\theta}{2\pi}.$$

The function s_h is twice differentiable in t with derivatives at 0 given by

$$s_h'(0) = \int_0^{2\pi} 2\Re \frac{\partial \Gamma}{\partial z}(e^{i\theta}, f^*(e^{i\theta}))^t h(e^{i\theta})\frac{d\theta}{2\pi}$$

and

$$s_h''(0) = \int_0^{2\pi} \left[\overline{h(e^{i\theta})}^t \frac{\partial^2 \Gamma}{\partial z \partial \bar{z}}(e^{i\theta}, f^*(e^{i\theta}))h(e^{i\theta}) + \Re \left(h(e^{i\theta})^t \frac{\partial^2 \Gamma}{\partial z^2}(e^{i\theta}, f^*(e^{i\theta}))h(e^{i\theta}) \right) \right] \frac{d\theta}{2\pi}$$

If f^* is a local directional optimizer for OPT^1, then for every $h \in A_N$, $s_h'(0) = 0$ and $s_h''(0) \geq 0$. It is now clear that II holds, so we now prove I. Since the real part of the complex valued linear functional Λ defined by

$$\Lambda(h) = \int_0^{2\pi} \frac{\partial \Gamma}{\partial z}(e^{i\theta}, f^*(e^{i\theta}))^t h(e^{i\theta})\frac{d\theta}{2\pi}$$

is 0 for every $h \in A_N$, it follows that Λ is the null functional. Then its kernel $\frac{\partial \Gamma}{\partial z}(\cdot, f^*(\cdot))$ must belong to the annihilator of A_N, this is the space χH^1_N. This proves one direction. The converse follows easily from the second order Taylor expansion of s_h about $t = 0$. ∎

Note that if for each $e^{i\theta}$, the function $\Gamma(e^{i\theta}, z)$ is strongly convex in z at $z_0 = f^*(e^{i\theta})$, i.e.

$$\overline{z}^t \, \frac{\partial^2 \Gamma}{\partial z \partial \overline{z}} \, (e^{i\theta}, f^*(e^{i\theta}))z \; + \; \Re \left(z^t \, \frac{\partial^2 \Gamma}{\partial z^2} \, (e^{i\theta}, f^*(e^{i\theta}))z \right) \geq 0, \qquad (\forall z) \tag{3}$$

then condition II is automatically satisfied. If in addition, (3) is satisfied with strict inequality at some θ_0 then condition I is sufficient for f^* to be an optimizer.

For a given measurable function $a \colon \partial D \to C^N$, set

$$\mathcal{N}(a) = \{ h \in H_N^2 \colon a^t(e^{i\theta})h(e^{i\theta}) = 0, \qquad \text{a.a.} e^{i\theta} \; \}$$

Theorem 2 below, which can be found in [HMer], gives optimality conditions for OPT^∞.

Theorem 2 *Let E be either H_N^∞ or A_N, and let Γ be of class C^3. Let $f^* \in E$ be such that $\frac{\partial \Gamma}{\partial z}(e^{i\theta}, f^*(e^{i\theta}))$ is bounded away from 0 on ∂D. If f^* is a local directional optimizer, then*

I a. The function $\Gamma(\,\cdot\,, f^*(\,\cdot\,))$ is constant on ∂D.

I b. There exist functions $F \in H_N^1$, and $\lambda \colon \partial \to R^+$ measurable and positive almost everywhere on ∂D such that

$$\lambda^{-1} \frac{\partial \Gamma}{\partial z} \, (\cdot, f^*) = \chi \, F$$

II. For every $h \in E \cap \mathcal{N}\left(\frac{\partial \Gamma}{\partial z}(\,\cdot\,, f^*)\right)$, $h \neq 0$,

$$\sup_\theta \left\{ \overline{h}^t \, \frac{\partial^2 \Gamma}{\partial z \partial \overline{z}} \, (\cdot, f^*)h + \Re \left(h^t \, \frac{\partial^2 \Gamma}{\partial z^2} \, (\cdot, f^*)h \right) \right\} \geq 0 \tag{4}$$

Conversely, conditions I *and* II *with strict inequality imply that f^* is a strict local directional optimizer.*

We point to the reader that both Theorems 1 and 2 hold true (with nearly identical proofs, see [HMer]) if the spaces A_N, H_N^∞, H_N^1 are replaced by their "real on the real axis" counterparts RA_N, RH_N^∞, RH_N^1, and if Γ is assumed to be real symmetric, i.e.

$$\Gamma(e^{i\theta}, z) = \Gamma(\overline{e^{i\theta}}, \overline{z}), \qquad (\forall e^{i\theta}), (\forall z)$$

Observe that in Theorem 2, if there exists θ_0 such that the function $\Gamma(e^{i\theta_0}, z)$ satisfies (3) at $z = f^*(e^{i\theta_0})$, then condition II. is automatically satisfied. Moreover, if (3) holds with strict inequality at some $\theta = \theta_0$, then conditions I.a and I.b are sufficient for optimality.

2.2 Connection between $OPT^\infty$and OPT^1

The similarity between condition I of Theorem 1 and condition Ib of Theorem 2 is striking. We now use this similarity to show that solutions to OPT^∞problems are also solutions to suitable OPT^1 problems. This suggests that numerical procedures for solving OPT^∞ based on OPT^1 could be effective.

Proposition 1 *let* Γ_1 *be of class* C^3*, and set* $E = A_N$*. Let* $f^* \in A_N$ *be such that* $\Gamma_1(e^{i\theta}, f^*(e^{i\theta})) \neq 0, \forall e^{i\theta}$*. If* f^* *is a local directional solution to* OPT^1*, then* f^* *is a local directional solution to* OPT^∞*for the function* $\Gamma(e^{i\theta}, z) = w(e^{i\theta})\Gamma_1(e^{i\theta}, z)$*, where* $w(e^{i\theta}) = 1/\Gamma_1(e^{i\theta}, f^*(e^{i\theta}))$*.*

Proof: If f^* is an optimizer to OPT^1 for Γ_1 but not to $OPT^\infty$for Γ , then there exists an $h \in A_n$ and $t > 0$ can be found arbitrarily small such that

$$\sup_{e^{i\theta}} \Gamma(e^{i\theta}, f^*(e^{i\theta})) > \sup_{e^{i\theta}} \Gamma(e^{i\theta}, f^*(e^{i\theta}) + th(e^{i\theta})) \tag{5}$$

But $\Gamma(e^{i\theta}, f^*(e^{i\theta}))$ is constant in θ and equal to 1, by the definition of Γ. Hence, the inequality 5 can be written as

$$1 > \sup_{\theta} \frac{1}{\Gamma_1(\cdot, f^*)} \Gamma_1(\cdot, f^* + t\,h)$$

i.e.,

$$\Gamma_1(e^{i\theta}, f^*(e^{i\theta})) > \Gamma_1(e^{i\theta}, f^*(e^{i\theta}) + th(e^{i\theta})) \qquad (\forall e^{i\theta})$$

Then

$$\int_0^{2\pi} \Gamma_1(e^{i\theta}, f^*(e^{i\theta}))\frac{d\theta}{2\pi} > \int_0^{2\pi} \Gamma_1(e^{i\theta}, f^*(e^{i\theta}) + th(e^{i\theta}))\frac{d\theta}{2\pi}$$

but this contradicts the hypothesis. ■

The following is a partial converse to Proposition 1. Its proof is simple and not given here.

Proposition 2 *Let* Γ *and* E *be as in Proposition 1. Let* f^* *be an optimizer to* OPT^∞ *and let* λ *be as in Thm 2.2 (I.b). Set* $\Gamma_1(e^{i\theta}, z) = \frac{1}{\lambda(e^{i\theta})}\Gamma(e^{i\theta}, z)$*. Then I of Theorem 2.1 holds for the function* Γ_1*. If in addition the function* $\Gamma(e^{i\theta}, z)$ *is convex in* z*, then* f^* *is a local directional optimizer for* OPT^1 *.*

3 The Newton iteration algorithm for solving OPT^{∞}

This section presents an algorithm for finding a function f^* that satisfies Ia. and Ib. of Theorem 2.1, when both f^* and Γ are smooth. It is shown that it is locally quadratically convergent provided technical hypotheses are satisfied.

3.1 Statement of the problem

We now take the equations in Ia. and Ib. of Theorem 2.1. and form a system of equations with them. This is stated below as Problem A.

Problem A. Find $(f^*, \lambda^*, F^*, \gamma^*)$ with $f^* \in H_N^{\infty}$, $\gamma^* \in R$, $F^* \in H_N^1$ and $\lambda^* : T \to R^+$ measurable and nonzero a.e. such that $(f, \lambda, F, \gamma) = (f^*, \lambda^*, F^*, \gamma^*)$ satisfies

$$\begin{aligned} \Gamma(\cdot, f) &= \gamma \\ \lambda^{-1} \tfrac{\partial \Gamma}{\partial z}(\cdot, f) &= \chi F \end{aligned} \tag{6}$$

The difficulty with applying Newton's method to equation (6) is that in a sense it has more unknown functions (four) than constraints. This means that local solutions are not isolated, rather there is a whole continuum of them. Newton's method is well known for its bad behavior in such circumstances and (before finding the method presented here) we experimented at length with Newton type solutions of (6) all of which performed terribly. Rather than working on Problem A directly, we will modify it first to obtain a problem with only two unknowns. This is the key step in our method to solving Problem A.

We proceed with formal calculations for a moment. Begin by eliminating indeterminacy in (6) by assuming $\lambda^{-1} \in L_1^2$ is such that $\int_0^{2\pi} \lambda(e^{i\theta})^{-1} \frac{d\theta}{2\pi} = 1$. Note that $P_{H^{2\perp}}[\, \Gamma(\cdot, f)\,] = 0$ since $\Gamma(\cdot, f)$ is a constant function and $P_{H^{2\perp}_{N,0}}[\, \chi F\,] = 0$ (provided $F \in L_N^2$). Thus we obtain the system[1]

$$\begin{aligned} P_{H_1^{2\perp}}[\, \Gamma(\cdot, f)\,] &= 0 \\ P_{H^{2\perp}_{N,0}}[\, \lambda^{-1} \tfrac{\partial \Gamma}{\partial z}(\cdot, f)\,] &= 0 \end{aligned} \tag{7}$$

in two unknowns, λ and f. Because of the way λ is normalized, one can write

$$\lambda^{-1} = 1 + \Re(\chi \beta) \tag{8}$$

where β is a scalar valued function which is analytic on the disk. Now, substituting (8) into (7) gives

$$T \begin{pmatrix} f \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{9}$$

---

[1] Note that we require that $1/\lambda$ is in L_1^2 instead of L_1^1. This allows the use of projections onto subspaces of L^2, on which our analysis is based upon. Later we will restrict our study to smooth functions λ, so the projection onto subspaces of L_N^2 is well defined.

where in equation (9),

$$T\begin{pmatrix} f \\ \\ \beta \end{pmatrix} = \begin{pmatrix} P_{H^{2\perp}_{N,0}}\left[(1+2\Re(\chi\beta))\ \frac{\partial\Gamma}{\partial z}(\cdot,f)\right] \\ \\ P_{H^{2\perp}}\,[\,\Gamma(\cdot,f)\,] \end{pmatrix} \tag{10}$$

At this point the basic idea is clear: to solve Problem A, it is enough to solve equation (9). Thus we state

Problem B. Find (f^*,β^*) with $1+2\Re(\chi\beta^*)$ positive a.e. on ∂D such that

$$T\begin{pmatrix} f^* \\ \beta^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{11}$$

3.2 Newton iteration for solving operator equations

We now discuss the solution of operator equations (e.g. equation (11)) with the iterative method of Newton. In terms of T' we have the Newton's step for updating a given pair (f,β):

$$\begin{pmatrix} \tilde{f} \\ \tilde{\beta} \end{pmatrix} = \begin{pmatrix} f \\ \beta \end{pmatrix} - \ T'^{-1}_{f,\beta}T\begin{pmatrix} f \\ \beta \end{pmatrix} \tag{12}$$

We refer to (12) as *Newton iteration*, and to the repeated application of (12) as the *Newton algorithm*. A standard fact about Newton iteration is

Theorem 3 *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be normed spaces, and suppose the operator $T:\mathcal{B}_1\longrightarrow\mathcal{B}_2$ is two times continuously Frechet differentiable in a neighborhood V of $x^*\in\mathcal{B}_1$. Assume also that $T[x^*]=0$, and that T'_{x^*} has a bounded inverse ${T'}^{-1}_{x^*}$.*

Then there exist a neighborhood W of x^* in $\mathcal{B}_1$ and a constant $c>0$ such that for every $x\in W$ the linear operator T'_x has a bounded inverse, and $\tilde{x}:=x-{T'}^{-1}_xT[x]$ satisfies $||\,\tilde{x}-x^*\,||_{\mathcal{B}_1}\le c\,||\,x-x^*\,||^2_{\mathcal{B}_1}$

Proof: The differentiability of T implies that there exist a constant $c_0>0$ and a neighborhood U of the origin in $\mathcal{B}_1$ such that

$$||\,T[x+y]-T[x]-T'_x[y]\,||_{\mathcal{B}_2}\le c_0\,||\,y\,||^2_{\mathcal{B}_1},\qquad(\forall y\in U)\,(\forall x\in V) \tag{13}$$

Since T'_{x^*} has a bounded inverse, there is a neighborhood W of x^* in $\mathcal{B}_1$ and a constant $c_1>0$ such that for every $x\in W$, T'_x has a bounded inverse with $||{T'}^{-1}_x||_{\mathcal{L}(\mathcal{B}_1,\mathcal{B}_2)}\le c_1$. We may assume $W\subset x^*+U$ without loss of generality. Take $y=x^*-x$ in (13) to obtain

$$||\,0-T[x]-T'_x[x^*-x]\,||_{\mathcal{B}_2}\le c_0\,||\,x^*-x\,||^2_{\mathcal{B}_1}, \tag{14}$$

Then

$$\begin{aligned} ||\,\tilde{x}-x^*\,||_{\mathcal{B}_1} &= ||\,x-{T'}^{-1}_xT[x]-x^*\,||_{\mathcal{B}_1} \\ &= ||\,{T'}^{-1}_x[T[x]-T'_x[x-x^*]]\,||_{\mathcal{B}_1} \\ &\le ||\,{T'}^{-1}_x\,||_{\mathcal{L}(\mathcal{B}_1,\mathcal{B}_2)}\,||\,T[x]-T'_x[x-x^*]\,||_{\mathcal{B}_2} \\ &\le c_0\,c_1\,||\,x^*-x\,||^2_{\mathcal{B}_1} \end{aligned} \tag{15}$$

■

One difficulty that arises with the definition of T as an operator on $H_N^\infty \times H_1^2$ is that the projection operators onto subspaces of L_N^2 are not bounded in the norm of the supremum. The space H_{N+1}^2 is not an option as domain for T either, since T is not defined at some points in this space. We shall see however that this problem can be overcome by considering T as an operator on a space of smooth functions. This will yield solutions to Problem B with Newton algorithm when the iteration is begun near a solution (f^*, β^*) that is smooth.

We now introduce the Sobolev spaces $\mathcal{S}_p$ of functions in L_{N+1}^2 and finite norm

$$|| f ||_{\mathcal{S}_p}^2 = ||\hat{f}_0||_{C^{N+1}}^2 + \sum_{n=-\infty}^{\infty} | n |^p || \hat{f}_n ||_{C^{N+1}}^2 \tag{16}$$

We note that the spaces $\mathcal{S}_p$ are Hilbert spaces. Also the projections $P_{H_{N+1}^2}$, $P_{H_{N+1,0}^2}$, $P_{H_{N+1}^{2\perp}}$, $P_{H_{N+1,0}^{2\perp}}$ are well defined on $\mathcal{S}_p$ and the relation

$$|| P_{H_{N+1}^2}[f] ||_{\mathcal{S}_p}^2 \; = \; || \hat{f}_0 ||_{C^{N+1}}^2 + \sum_{n=1}^{\infty} |n|^p || \hat{f}_n ||_{C^{N+1}}^2 \; \leq \; || f ||_{\mathcal{S}_p}^2 \tag{17}$$

illustrates that their norm is no larger than 1.

We also need the set $\mathcal{A}_{N+1}^\infty = (\cap_{n=0}^\infty \mathcal{S}_p) \cap H_{N+1}^2$. For each fixed $p > 0$, the space $\mathcal{A}_{N+1}^\infty$ is a normed (but not complete) space[2] when equipped with the Sobolev norm $|| \cdot ||_{\mathcal{S}_p}$. The set of all $\overline{g}$ where $g \in \mathcal{A}_{N+1}^\infty$ is denoted $\overline{\mathcal{A}_{N+1}^\infty}$.

Lemma 1 *The linear operator in $\mathcal{S}_p$ of multiplication by a matrix valued function with entries in $\mathcal{A}_1^\infty$ is bounded.*

Proposition 3 *Let Γ be a smooth function of $e^{i\theta}$ and z. Let $p \geq 0$ and equip $\mathcal{A}_{N+1}^\infty$ with the norm of $\mathcal{S}_p$. Then the relation (10) defines an operator $T : \mathcal{A}_{N+1}^\infty \to \overline{\mathcal{A}_{N+1}^\infty} \cap (H_{N,0}^{2\perp} \times H_1^{2\perp})$. Moreover, T is smooth at every (f, β) in $\mathcal{A}_{N+1}^\infty$.*

We don't prove Lemma 1 or Proposition 3 but merely remark that the operator $T'_{f,\beta}$ is calculated explicitly below in terms of Γ, and that similar calculations can be carried out to obtain $T_{f,\beta}^{(n)}$ explicitly for any n. From these calculations and the Taylor expansion of T with integral remainder follows that T is smooth on $\mathcal{A}_{N+1}^\infty$ in the norm of $\mathcal{S}_p$.

We now derive the formula for the Jacobian $T'_{f,\beta}$ of T, which we obtain by linearizing (10). The following expansions will be needed.

$$\begin{array}{rcl} \Gamma(\cdot, f+\epsilon) & = & g + 2\Re(a^t \epsilon) + \overline{\epsilon}^t A \epsilon + \Re(\epsilon^t B \epsilon) + \cdots \\ \frac{\partial \Gamma}{\partial z}(\cdot, f+\epsilon) & = & a + A\overline{\epsilon} + B\epsilon + \cdots \end{array} \tag{18}$$

---

[2] The vector space $\mathcal{A}_{N+1}^\infty$ can be given a structure of Frechet space through the use of the family of norms $|| \cdot ||_{\mathcal{S}_p}$ to define a locally convex topology. The difficulty here is that $\mathcal{A}_{N+1}^\infty$ with such topology is not normable, and in this case one cannot take advantage of Theorem 3 or a variation of it to obtain an algorithm.

where

$$g = \Gamma(\cdot, f), \quad a = \tfrac{\partial \Gamma}{\partial z}(\cdot, f), \quad A = \tfrac{\partial^2 \Gamma}{\partial \overline{z} \partial z}(\cdot, f), \quad \textit{and} \quad B = \tfrac{\partial^2 \Gamma}{\partial z^2}(\cdot, f). \tag{19}$$

Consider the expression

$$T \begin{pmatrix} f + \epsilon \\ \beta + \delta \end{pmatrix} = \begin{pmatrix} P_{H^{2\perp}_{N,0}} \left[\, (1 + 2\Re(\chi(\beta + \delta))) \, \tfrac{\partial \Gamma}{\partial z}(\cdot, f + \epsilon) \, \right] \\ P_{H^{2\perp}_1} \left[\, \Gamma(\cdot, f + \epsilon) \, \right] \end{pmatrix} \tag{20}$$

Substituting (18) in (20) and dropping all terms which are not linear or constant in δ or ϵ yields

$$\begin{aligned} T \begin{pmatrix} f \\ \beta \end{pmatrix} \quad + T'_{f,\beta} \begin{pmatrix} \epsilon \\ \delta \end{pmatrix} &= \\ &\begin{pmatrix} P_{H^{2\perp}_{N,0}} [\, (1 + 2\Re(\chi(\beta + \delta))) \, (a + A\overline{\epsilon} + B\epsilon) + 2\Re(\chi\delta) a \,] \\ P_{H^{2\perp}_1} [\, g + 2\Re a^t \epsilon \,] \end{pmatrix} \end{aligned} \tag{21}$$

Thus we have that the Jacobian of T is given by

$$T'_{f,\beta} \begin{pmatrix} \epsilon \\ \delta \end{pmatrix} = \begin{pmatrix} P_{H^{2\perp}_{N,0}} [\, (1 + 2\Re(\chi\beta)) \, (A\overline{\epsilon} + B\epsilon) + 2a\Re(\chi\delta) \,] \\ P_{H^{2\perp}_1} [\, 2\Re a^t \epsilon \,] \end{pmatrix} \tag{22}$$

An elegant way to write T' is

Lemma 2 *The operator $T'_{f,\beta}$ in (22) can be written in the form*

$$T'_{f,\beta} = \begin{pmatrix} I_N & 0 \\ 0 & \overline{\chi} \end{pmatrix} \mathcal{C} \mathcal{T}_{M_1} + \begin{pmatrix} \chi I_N & 0 \\ 0 & 1 \end{pmatrix} \mathcal{H}_{M_2} \tag{23}$$

where $\mathcal{T}_{M_1}$ is a Toeplitz operator with symbol $M_1 = \begin{pmatrix} wA & \chi\overline{a} \\ \overline{\chi} a^t & 0 \end{pmatrix}$, $\mathcal{H}_{M_2}$ is the Hankel operator with symbol $M_2 = \begin{pmatrix} \overline{\chi} w B & a \\ a^t & 0 \end{pmatrix}$, $\mathcal{C}$ is the operator $f \rightarrow \overline{f}$, I_N is the $N \times N$ identity matrix, and $w = 1 + 2\Re\chi\beta$.

A proof of Proposition 2 follows from the relations $P_{H^{2\perp}_{N+1,0}}[\overline{M}\overline{\eta}] = \mathcal{C} P_{H^2_{N+1}}[M\eta]$ and $\overline{\chi} P_{H^{2\perp}_{N+1,0}}[\chi M \eta] = P_{H^{2\perp}_{N+1}}[M\eta]$ applied to (22).

One can see in representation (23) that $T'_{f,\beta}$ has the form of a " Conjugate of Toeplitz plus Hankel" operator (except for a small change in the range, which will turn out to be unimportant for our analysis of $T'_{f,\beta}$; see Subsection 3.4). One consequence of this as pointed out to us by Ali Sayed is that $T'_{f,\beta}$ can be inverted numerically with fast algorithms similar to those described in [SLK]. This enhances the practical appeal of Newton's method applied to the operator T.

An important fact is that the operator T as given in Proposition 3 is not defined on a Banach space, thus preventing a direct application of the general result Theorem 3. Analysis of the proof of the latter reveals that completeness is needed only to ensure the existence of a neighborhood W of the root x^* of $T[x] = 0$, such that for every $x \in W$ the operator T'_x is invertible. Thus the conclusion of Theorem 3 may still hold provided it is possible to use the particular structure in our situation to find such a set W. This indeed is the case, as we shall see in Theorem 6.

Suppose that $(f^*, \beta^*) \in \mathcal{A}^\infty_{N+1}$ is a solution to Problem B. We claim that *under technical hypotheses and for any* $p \geq 0$, *Newton algorithm applied to the operator* T *is locally second order convergent in* $\mathcal{A}^\infty_{N+1}$ *(equipped with the norm of* $\mathcal{S}_p$*) to* (f^*, β^*) (see Corollary 1 in Subsection 5). For the proof of this claim we need, in addition to the smoothness of T, the existence of a neighborhood of (f^*, β^*) on which $T'_{f,\beta}$ is invertible.

While the claim is vaguely stated, the next section lays the groundwork which allows us to state its hypotheses precisely and motivate them.

3.3 Background

A list of results that will be needed later is stated here.

Definition 2 *A function* $F \in H^1_N$ *is called*

i. outer if the only scalar inner functions q *such that* F *belongs to* qH^1_N *are the constant functions.*

ii. strong outer if it is outer and for every $e^{i\phi} \in \partial D$ *the function* $(\chi - e^{i\phi})^{-2}F$ *is not in* H^1_N.

iii. unitary if $||F(e^{i\theta})||_{C^N} = 1$ *for almost all* $e^{i\theta}$.

Proposition 4 below is an observation in [C].

Proposition 4 *If* $N > 1$ *and* $F \in H^1_N$, *then there exist* $E \in H^1_N$ *outer unitary and* $f \in H^1_1$ *such that* $F = fE$. *Moreover, if* $F_1 \in H^1_N$ *is such that* $F_1 = gE$ *for some measurable scalar function* g, *then* $g \in H^1_1$.

The following result is a refinement of Ib. in Theorem 2 (see [HMer2]).

Proposition 5 *Let* Γ *and* $f^* \in H^\infty_N$ *be as in Theorem 2, and suppose that* f^* *satisfies Ia. and Ib. of Theorem 2. Then there exist a strong outer function* $F_0 \in H^1_N$, *a number* $n \geq 1$ *and a positive measurable function* λ *such that the function* $a(e^{i\theta}) = \frac{\partial\Gamma}{\partial z}(e^{i\theta}, f^*(e^{i\theta}))$ *satisfies*

$$a = \lambda \chi^n F_0 \tag{24}$$

The functions λ *and* F_0 *in (24) are uniquely determined if* $||1/\lambda||_{L^1} = 1$.

Proposition 6 *Under the hypothesis and notation of Proposition 5, there exists a* $N \times (N-1)$*-matrix valued function* Θ *with entries in* H^∞_1 *such that* $\Theta(e^{i\theta}) : C^{N-1} \to C^N$ *is an isometry for almost all* $e^{i\theta}$, *and*

$$\mathcal{N}(a) = \{h \in H^2_N : \ a(e^{i\theta})^t h(e^{i\theta}) = 0 \ a.a. e^{i\theta} \ \} = \Theta H^2_{N-1} \tag{25}$$

It was shown in [HMer2] that one can obtain a parameterization of $\mathcal{N}(a)$ similar to (25) but with Θ with continuous entries if a is of class C^1. The proof offered below follows an argument in [C, p.292].

Proof: It is clear that for almost all $e^{i\theta}$, the subspace of C^N of vectors v such that $a(e^{i\theta})^t v = 0$ is $N-1$ dimensional. The space $\mathcal{N}(a)$ is a closed, simply invariant subspace of H^2_N with trivial remote past. Thus by Theorem 1 in [C] there exists Θ, such that $\Theta(e^{i\theta}) : C^{N-1} \to C^N$ is an isometry for almost all $e^{i\theta}$, and such that $\mathcal{N}(a) = \Theta H^2_{N-1}$. ∎

The number n in (24) has been proved in [Mer] to be equal to 1 in generic scalar cases (i.e., $N = 1$). The same statement for $N > 1$ has not been proved but we conjecture this to hold there too.

Conjecture 1 *For generic Γ, the number n in Proposition 5 is 1.*

Lemma 3 *If $\Psi \in H^1_1$ is such that $\Re\, e^{i\theta}\Psi(e^{i\theta}) \geq 0$ for almost all $e^{i\theta} \in \partial D$, then $\Psi = 0$.*

Proof. The function $f(z) = z\Psi(z)$ is analytic on the unit disk D, and maps D onto a neighborhood of the origin if $\Psi \neq 0$. By the Poisson integral representation of f, the inequality $\Re\, f(e^{i\theta}) \geq 0$ implies that $f = 0$, thus $\Psi = 0$. ∎

Lemma 4 *Let M be a $m \times m$ matrix valued function on the unit circle with entries in $\mathcal{A}^\infty_m$. If the function M takes self-adjoint, positive definite values and is such that $det(M)$ is not 0 on ∂D, then the Toeplitz operator $\mathcal{T}_M : f \to P_{H^2_{N+1}}[f]$ is an invertible map from $\mathcal{S}_p \cap H^2_{N+1}$ to $\mathcal{S}_p \cap H^2_{N+1}$ for any $p \geq 0$.*

Proof: The outer spectral factors v and v^* of M satisfy $v^*v = M$ and are smooth and uniformly invertible since M is smooth and uniformly invertible. The inverse is

$$\mathcal{T}_M^{-1} = P_{H^{2\perp}_m} v^{*-1} P_{H^{2\perp}_m} v^{-1} P_{H^{2\perp}_m} \tag{26}$$

which is a composition of bounded operators on $\mathcal{S}_p \cap H^2_{N+1}$ for any p. ∎

3.4 T' has Fredholm index 0

When (f, β) is smooth, the operator $T'_{f,\beta}$ can be defined on a Banach space containing $\mathcal{A}^\infty_{N+1}$. This will allow the use of the theory of Fredholm operators.

The following Lemma 5 is a consequence of (23), Lemma 1 and Lemma 4.

Lemma 5 *If (f, β) is in $\mathcal{A}^\infty_{N+1}$, then $T'_{f,\beta}$ can be extended to a bounded operator from $\mathcal{S}_p \cap H^2_{N+1}$ to $\mathcal{S}_p \cap (H^{2\perp}_{N,0} \times H^{2\perp}_1)$.*

If $\mathcal{B}_1$ and $\mathcal{B}_2$ are Banach spaces, we denote with $\Phi(\mathcal{B}_1, \mathcal{B}_2)$ the collection of Fredholm operators in $\mathcal{L}(\mathcal{B}_1, \mathcal{B}_2)$ (See [BS,page 17] for definitions and results).

Theorem 4 *With the notation used in (19) and in Lemma 2, suppose that $(f, \beta) \in \mathcal{A}^\infty_{N+1}$ is such that A is positive definite, w is positive*[3] *and that $det(A)$ and a are bounded away from 0. Then the operator $T'_{f,\beta}$ is in $\Phi(\mathcal{S}_p \cap H^2_{N+1}, \mathcal{S}_p \cap (H^{2\perp}_{N,0} \times H^{2\perp}_1))$. Moreover, the Fredholm index of $T'_{f,\beta}$ is 0.*

---

[3] Addendum to published version: A hypothesis has been dropped. We must also assume that the integer n in (28) is less than or equal to 1. The problem is that Lemma 4 does not apply to $\mathcal{T}_{M_1}$ directly since M_1 is not positive definite. Fortunately, this error has no impact on the rest of the paper because for our main theorems 5 and 6 we assume n=1. The proof of the corrected theorem which we just stated relies on the special structure of M_1. It can be gotten directly from the arguments in the next section which use this structure thoroughly.

Proof: Lemma 4 implies that $\mathcal{T}_{M_1}$ is invertible as an operator from $\mathcal{S}_p \cap H^2_{N+1}$ to $\mathcal{S}_p \cap (H^{2\perp}_{N,0} \times H^{2\perp}_1)$. Since $\mathcal{H}_{M_2}$ is a compact operator on $\mathcal{S}_p \cap H^2_{N+1}$ with range in $\mathcal{S}_p \cap (H^{2\perp}_{N,0} \times H^{2\perp}_1)$ (since the symbol is smooth), the conclusion follows from these facts and representation (23). ■

An immediate corollary to Theorem 4 is that $T'_{f,\beta}$ as an operator on $\mathcal{S}_p \cap H^2_{N+1}$ has closed range in $\mathcal{S}_p \cap H^{2\perp}_{N,0}$ with codimension equal to the dimension of the null space of $T'_{f,\beta}$.

3.5 T' has trivial null space

We now restrict attention to H^2_{N+1} and show that under certain technical hypotheses and if (f^*, β^*) is a smooth solution to Problem B, T'_{f^*,β^*} as an operator on H^2_{N+1} has no null vector in it. This proves that T'_{f^*,β^*} has no null vector in $\mathcal{S}_p \cap H^2_{N+1}$ for all p, since $\mathcal{S}_p \subset \mathcal{S}_0$. On the other hand, the last paragraph in the previous subsection shows that in this case the map $T'{}_{f^*,\beta^*} : \mathcal{S}_p \cap H^2_{N+1} \to \mathcal{S}_p \cap (H^{2\perp}_{N,0} \times H^{2\perp}_1)$ is invertible for all p. We begin with a lemma.

Lemma 6 *Let A_0, B_0, be $m \times m$ matrix valued functions with entries in L^2_1, and such that A_0 takes self-adjoint, positive definite values and it has $det(A_0)$ bounded away from 0 on T. Then there exist a $m \times m$ matrix valued analytic function μ with entries in H^2_1, with analytic inverse μ^{-1} of the same type, so that the equation*

$$P_{H^{2\perp}_{m,0}}\left(A_0\overline{\phi} + B_0\phi\right) = 0 \tag{27}$$

has a nontrivial solution $\phi \in H^2_m$ if and only if there exist nontrivial solutions to the equation

$$\mathcal{H}_M(\psi) = \overline{\chi}\overline{\psi} \tag{28}$$

where the symbol M of the Hankel operator $\mathcal{H}_M$ is given by

$$M = -\overline{\chi}(\mu^{-1})^t B_0 \mu^{-1} \tag{29}$$

In fact, μ satisfies

$$A_0 = \mu^t \overline{\mu} \tag{30}$$

i.e., μ is the outer spectral factor of A_0.

Proof: The hypotheses on A_0 guarantee the existence of a spectral factor μ satisfying (30), see for example [Hs]. We proceed to rewrite (27) in a way that will make obvious the conclusion of the theorem. Pick $G \in H^2_m$ such that

$$A_0\overline{\phi} + B_0\phi = \chi G \tag{31}$$

Hence

$$\overline{\mu}\overline{\phi} + (\mu^{-1})^t B_0 \mu^{-1}\mu\phi = (\mu^{-1})^t \chi G \tag{32}$$

Set $\psi = \mu\phi$ and apply the projection onto $H^{2\perp}_m$ to both members of (32) to obtain

$$\overline{\chi}\overline{\psi} + P_{H^{2\perp}_m}[\overline{\chi}(\mu^{-1})^t B_0 \mu^{-1}\psi] = 0 \tag{33}$$

Finally note that equation (33) can be written as

$$P_{H^{2\perp}_m}[M\psi] = \overline{\chi}\overline{\psi} \tag{34}$$

■

Theorem 5 *Let* $(f^*, \beta^*) \in \mathcal{A}_{N+1}^{\infty}$ *and set* $g = \Gamma(\cdot, f^*)$, $a = \frac{\partial \Gamma}{\partial z}(\cdot, f^*)$, $A = \frac{\partial^2 \Gamma}{\partial \overline{z} \partial z}(\cdot, f^*)$, *and* $B = \frac{\partial^2 \Gamma}{\partial z^2}(\cdot, f^*)$. *Suppose that* $T[(f^*, \beta^*)^t] = (0,0)^t$ *and that*

i. *Equation (24) holds with* $n = 1$ *for some functions* $\lambda = 1/w^*$ *and* F_0.

ii. *If* Θ *is the function whose existence is guaranteed by Proposition 6, then the function* $A_0 = w^* \Theta^t A \overline{\Theta}$ *has* $det(A_0)$ *bounded away from* 0 *on* T.

Then the operator equation

$$T'_{(f^*, \beta^*)} \begin{pmatrix} \epsilon \\ \delta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{35}$$

has $(\epsilon, \delta) = (0,0)$ *as the only solution in* $H_1^2 \times H_N^2$ *if and only if the equation*

$$\mathcal{H}_M(\psi) = \overline{\chi} \overline{\psi} \tag{36}$$

has only the trivial solution. Here the Hankel operator $\mathcal{H}_M$ *has symbol* $M = -\overline{\chi}(\mu^{-1})^t w^* \Theta^t B \Theta \mu^{-1}$, *where* μ *is the outer spectral factor of* A_0 *given in Theorem 6.*

Proof: Consider (ϵ, δ) that satisfy (35), and suppose that equation (36) has only the the trivial solution. From the relation

$$P_{H^{2\perp}}(2\Re a \epsilon) = 0 \tag{37}$$

we see that $2\Re a(\cdot)\epsilon(\cdot)$ is a constant function of $e^{i\theta}$:

$$2\Re a(e^{i\theta})\epsilon(e^{i\theta}) = c_0 \qquad a.a.e^{i\theta} \tag{38}$$

We now show that $c_0 = 0$. Since $T(f^*, \beta^*)^t = (0,0)^t$, there exists $F \in H_N^2$ such that

$$(1 + 2\Re\chi\beta^*)a = \chi F \tag{39}$$

Combine (38) with (39) to obtain

$$2\Re\chi F^t \epsilon = c_0(1 + \Re\chi\beta^*)^{-1} \tag{40}$$

Since $1 + \Re\chi\beta^*$ does not change sign on the unit circle, Lemma 1 implies that

$$F(e^{i\theta})^t \epsilon(e^{i\theta}) = 0, \qquad a.a.e^{i\theta} \tag{41}$$

thus $c_0 = 0$ and we have that

$$a(e^{i\theta})^t \epsilon(e^{i\theta}) = 0, \qquad a.a.e^{i\theta} \tag{42}$$

We now proceed to show that $\epsilon = 0$. This is obvious when $N = 1$, so we assume $N > 1$ for now. Set $w^* = 1 + \Re(\chi\beta^*)$. Since (ϵ, δ) satisfies (35), there exists a function $G \in H_N^2$ such that

$$w^*(A\overline{\epsilon} + B\epsilon) + 2\Re(\chi\delta)a = \chi G \tag{43}$$

Write $\epsilon = \Theta\tilde{\epsilon}$ where $\tilde{\epsilon} \in H_{N-1}^2$ and multiply (43) on the left by Θ^t to obtain

$$w^*\Theta^t A \overline{\Theta\tilde{\epsilon}} + w^*\Theta B \Theta\tilde{\epsilon} = \chi\Theta^t G \tag{44}$$

Equation 44 leads to

$$P_{H^{2\perp}_{N-1,0}}\left(w^*\Theta^t A\overline{\Theta\tilde{\epsilon}} + w^*\Theta^t B\Theta\tilde{\epsilon}\right) = 0 \tag{45}$$

Lemma 6 can be applied to (45) to conclude that either $\tilde{\epsilon} = 0$ or there are nontrivial solutions to equation (36). However, the latter is not possible by assumption. Thus $\epsilon = \Theta\tilde{\epsilon} = 0$.

We next prove that $\delta = 0$. The relation $\epsilon = 0$ reduces (35) to the equation

$$P_{H^{2\perp}_{N,0}}(2\Re(\chi\delta)a) = 0 \tag{46}$$

Equation (46) can be written as

$$\chi G_2 = 2\Re(\chi\delta)a \tag{47}$$

where G_2 is a function in H^2_N. By the representation of a in (24) we have that

$$G_2 = 2\Re(\chi\delta)w^* F_0 \tag{48}$$

The case $N = 1$ is treated first. Since F_0 is outer, one can pick analytic polynomials p_k, $k = 1, 2, \ldots$ such that $p_k F_0 \to 1$ as $k \to \infty$ in H^1_1 ([dLR], page 474. See also [C], page 294). Since the function $2\Re(\chi\delta)w^*$ belongs to L^1_1, then $p_k G_2 = 2\Re(\chi\delta)w^* p_k F_0 \to 2\Re(\chi\delta)w^*$ as $k \to \infty$ and we have that $2\Re(\chi\delta)w^*$ is in H^1_1. But the latter function is real valued, so it must be the constant function. Now $\delta \neq 0$ is impossible for in that case the real valued function $2\Re(\chi\delta)w^*$ must change sign by Lemma 3. Thus $\delta = 0$.

Now consider $N > 1$. Proposition 4 applied to (48) guarantees the existence of scalar functions $f, g \in H^1_1$ and an outer unitary function $E \in H^1_N$ such that

$$gE = 2\Re(\chi\delta)w^* fE \tag{49}$$

Observe that the factorization $F_0 = fE$ guarantees f to be a scalar outer function. Equation (49) can be rewritten as

$$g = 2\Re(\chi\delta)w^* f \tag{50}$$

Hence we are back in the situation treated in the scalar case $N = 1$, and the conclusion again is that $\delta = 0$.

So far half of the Theorem has been proved. The other half follows from the observation that if equation (36) has nontrivial solutions, then one can reverse the steps leading to (45) to deduce that there are nontrivial solutions to (35). ■

3.6 Derivation of the main convergence result

Theorem 5 in Subsection 3.5 above says that T'_{f^*,β^*} is invertible as an operator on Sobolev space when (f^*, β^*) is a smooth solution to Problem B. The next result shows that $T'_{f,\beta}$ is actually invertible for all (f, β) in a neighborhood of (f^*, β^*). Moreover, its inverse maps smooth functions to smooth functions. A corollary of this and the smoothness of T is the basic Newton estimate that gives the local quadratic convergence of the algorithm.

Theorem 6 *Let $\Gamma(e^{i\theta}, z)$ be smooth and strictly plurisubharmonic in z. Let $(f^*, \beta^*) \in \mathcal{A}^{\infty}_{N+1}$ be a solution to Problem B, i.e., $T[(f^*, \beta^*)^t] = (0,0)^t$ and $w = 1 + 2\Re\chi\beta^* > 0$. Set $g = \Gamma(\cdot, f^*)$, $a = \frac{\partial\Gamma}{\partial z}(\cdot, f^*)$, $A = \frac{\partial^2\Gamma}{\partial\bar{z}\partial z}(\cdot, f^*)$, and $B = \frac{\partial^2\Gamma}{\partial z^2}(\cdot, f^*)$. Assume the following statements:*

- *i. The function a is not zero on ∂D.*
- *ii. The function $\overline{\chi}wa$ is outer analytic.*
- *iii. Equation (36) has only the trivial solution.*

Then there exists $W \subset \mathcal{A}^{\infty}_{N+1}$ such that for each $p \geq 0$ the set W is a neighborhood of (f^*, β^*) in the norm of $\mathcal{S}_p$, and for all $(f, \beta) \in W$, the operator $T'_{f,\beta} : \mathcal{A}^{\infty}_{N+1} \to \overline{\mathcal{A}^{\infty}_{N+1}} \cap (H^{2\perp}_{N,0} \times H^{2\perp}_{1})$ is invertible.

Before presenting the proof it is necessary to comment on the hypotheses to Theorem 6. The assumption on Γ being strictly p.s.h. (plurisubharmonic) in z can be eliminated but at the cost of having other technical hypotheses which are difficult to verify. All functions Γ that arise in applications are strictly p.s.h. Assumption i. is already found in the Theorem 2 characterizing solutions to OPT^{∞}, but here the statement is simpler due to the smoothness of f^*. Assumption ii. holds for generic Γ when $N = 1$, and was conjectured to hold also when $N > 1$ in Subsection 3.3 (see Proposition 5 and Conjecture 1). Again, smoothness permits a simple statement here. Assumption iii. requires in particular that the number 1 is not a singular value of certain Hankel operator. This is likely to hold for generic Γ, but we do not pursue this here to give a proof.
Proof of Theorem 6. The hypotheses of the theorem imply the hypotheses of Theorem 5, which guarantees that T'_{f^*,β^*} as an operator on H^2_{N+1} has trivial kernel. Since by Theorem 4 this operator is Fredholm, it is also invertible. By continuity of T'_x in x there exists an open set W in H^2_{N+1} such that $x \in W$ implies that T'_x is invertible as an operator with domain H^2_{N+1}. Therefore for any (f, β) in $\mathcal{A}^{\infty}_{N+1} \cap W$, the operator $T'_{f,\beta}$ has trivial kernel on $\mathcal{S}_p \cap H^2_{N+1}$ and it is invertible since it is Fredholm by Theorem 4.

Let (f, β) be an arbitrary element in $W \cap H^2_{N+1}$, and let $p \geq 0$. If h is in $\overline{\mathcal{A}^{\infty}_{N+1}} \cap (H^{2\perp}_{N,0} \times H^{2\perp}_{1})$ then $(T'_{f,\beta})^{-1}[\ h\ ]$ is in $\mathcal{S}_p \cap H^2_{N+1}$. But since p is arbitrary we have that necessarily $(T'_{f,\beta})^{-1}[\ h\ ]$ belongs to $\mathcal{A}_{N+1}$. ■

The following corollary gives the basic estimate for the Newton iteration.

Corollary 1 *Fix $p \geq 0$. There exists a constant $c > 0$ and a neighborhood $V \subset \mathcal{A}^{\infty}_{N+1}$ of (f^*, β^*) (in the norm of $\mathcal{S}_p$) such that for every $(f, \beta) \in V$ the Newton iteration (12) produces $(\tilde{f}, \tilde{\beta}) \in \mathcal{A}^{\infty}_{N+1}$ such that*

$$\left\| \begin{pmatrix} \tilde{f} - f^* \\ \tilde{\beta} - \beta^* \end{pmatrix} \right\|_{\mathcal{S}_p} \leq c \left\| \begin{pmatrix} f - f^* \\ \beta - \beta^* \end{pmatrix} \right\|^2_{\mathcal{S}_p} \tag{51}$$

4 Comparison of Newton iteration for $OPT^{\infty}$to the Disk iteration

In this section we present results of numerical calculations to solve OPT^{∞}in specific instances with the algorithm introduced in this paper. The aim is to illustrate the performance of an implementation of the algorithm, and the use of the optimality conditions in Theorem 2 to assess progress of the iteration through diagnostic derived directly from these optimality conditions.

The Newton Iteration method for solving OPT^{∞}introduced in this paper is unique in several respects. It does not rely on solving for the eigenvector of a Hankel operator associated with the largest singular value, as do methods such as those which iterate solutions to Nehari's problem (see [Yng],[HMer1]). We discuss this from a theoretical standpoint below in in Subsections 4.3and 4.5. The Newton iteration method is second order convergent, which to our knowledge is unique. This compares with other methods which are (in the vector valued case $N > 1$) typically linearly convergent. In order to show this striking difference, we consider a somewhat simple function Γ, and optimize it using two methods: Newton iteration, and Disk iteration [HMer1].

The software used to generate the numbers in this section it the program Anopt, developed in the Laboratory for Mathematics and Statistics at UCSD. It is available upon request[4]. Anopt is written in the language Mathematica and runs on any machine that supports Mathematica version 2.0 or higher.

4.1 Diagnostics

One practical consequence of the optimality conditions in Theorem 2 is that they lead to optimality tests, i.e., positive valued operators on A_N that are zero at a given function f when the optimality conditions Ia. and Ib. are satisfied. The value of these tests is that a discrete version of them can be easily implemented on the computer, and in practice can be used as an indirect measure of "closeness" to a solution to OPT^{∞}. For given $\Gamma(e^{i\theta}, z)$ and $f \in A_N$ set

$$a(e^{i\theta}) = (a_1(e^{i\theta}), \ldots, a_N(e^{i\theta})) := \frac{\partial \Gamma}{\partial z}(e^{i\theta}, f(e^{i\theta}))$$

and define OT_1, OT_2 as

$$\mathrm{OT}_1(f) := \frac{\sup_{e^{i\theta}} \Gamma(e^{i\theta}, f(e^{i\theta})) - \inf_{e^{i\theta}} \Gamma(e^{i\theta}, f(e^{i\theta}))}{\sup_{e^{i\theta}} \Gamma(e^{i\theta}, f(e^{i\theta}))} \tag{52}$$

$$\mathrm{OT}_2(f) := \begin{cases} Wind(a;0) - 1 & \text{if } N = 1 \\ \sum_{j=1}^{N} ||P_{H_N^{2\perp}}[a_j/a_1]||_{H_1^2} & \text{if } N > 1 \end{cases} \tag{53}$$

At the risk of distracting the reader we now verify that if the functions involved are well behaved then OT_1 and OT_2 are zero at functions f that satisfy the optimality conditions in Theorem 2.

It is obvious that $OT_1(f) = 0$ is equivalent to condition Ia in Theorem 2 being satisfied by f.

---

[4] Write electronic mail request to anopt@math.ucsd.edu

Suppose that a is of class C^1, and that a_1 is has no zeros on the circle. If $OT_2(f) = 0$, then one has that the function $(1/a_1)\ a$ is analytic, so $a = a_1 \times (analytic\ function)$, and it is clear by writing a_1 as a positive valued function times an analytic function that condition Ib of Theorem 2 is satisfied at f.

Now assume that f is a solution to OPT^∞, so in particular it satisfies condition Ib. If $N = 1$ then $Wind(a; 0) = 1$ for generic Γ [Mer], so $OT_2(f) = 0$ in such cases. For the case $N > 1$ assume first that $Wind(a_1; 0) = 1$. Then the equality $a_1 = \lambda\chi F_1$ from condition Ib and smoothness of a say that F_1 is outer. Hence $a_j/a_1 = F_j/F_1$ is analytic, so we have that $P_{H_1^{2\perp}}[a_j/a_1] = 0$. One last remark is that the formula for OT_2 may be modified so that a_1 is replaced by any coordinate function a_{j_0} for which $Wind(a_{j_0}; 0) = 1$. We now use an euristic argument to show why this latter condition is to be expected. A coordinate function a_j describes a curve in the plane that (generically) does not pass through 0. For j fixed, define the function $\Gamma_0(e^{i\theta}, z) = \Gamma(e^{i\theta}, f_1(e^{i\theta}), \ldots, z_j, \ldots, f_N(e^{i\theta}))$ of $e^{i\theta}$and the scalar variable z. The problem OPT^∞associated with Γ_0 has a local solution at f_j. But for generic Γ_0, one has $Wind(a_j; 0) = 1$ [Mer].

If $N = 1$ and a is zero at some $e^{i\theta}$, then $OT_2(f)$ is not defined. When $N > 1$, $OT_2(f)$ may make sense only if f is near a local solution f^*.

4.2 Newton iteration runs

We shall solve $OPT^\infty$when Γ is given by

$$\Gamma(e^{i\theta}, z_1, z_2) = |z_1|^2 + |z_2|^2 + |100 + e^{i\theta} z_1 + .1(z_1 + z_2 + z_1 z_2)|^2 + |100 + e^{i\theta} z_2 + .1(z_1 + z_2 + z_1 z_2)|^2 \quad (54)$$

The (absolute) optimal value is $\gamma^* = 3800$, attained at two different points in function space: the constant functions $f_1^* = (30, -30)$ and $f_2^* = (-30, 30)$. There are no other local solutions to OPT^∞. The saddle point value 20,000 occurs at the constant function $f_3^* = (0, 0)$ (here $OT_1(f_3^*) = 0$ and $OT_2(f_3^*) = 0$, but there is no local minimum there since condition II in Theorem 2 is not satisfied. For more details on this example see [HMer1]).

A final observation is that the Newton iteration method runs presented here do not include a linesearch along the direction given by one step in the method, i.e. a full step is taken each time. This is in contrast with the Disk iteration method (see [HM]), which depends on the linesearch in a strong manner.

The computer output shown in Tables 1 – 5 shows several columns. The meaning of each one is as follows:

it	Iteration number
Value	Current value of $\sup \Gamma(\cdot, f))$
|| f - f* ||	The true error (f* is the solution)
OT1	Value of OT_1 at current f.
OT2	Value of OT_2 at current f.
NED	A measure of numerical error.

The discretization of the problem is carried out by sampling functions on a grid of equally spaced points on the unit circle. In all the examples in Tables 1 – 5 the number of samples is 256.

Table 1

```
It :        Value        :||fk-f*||: Optimality Tests  :  Error
   :      SupGamma       :         :   OT1       OT2   :   NED
-----------------------------------------------------------------
 0 : 3.8165268771511E+03 : 5.7E-01 : 8.3E-03 : 1.6E+00 :   N/A
 1 : 3.8009177903975E+03 : 5.3E-02 : 4.9E-04 : 3.9E-01 : 8. E-13
 2 : 3.8000812563274E+03 : 2.7E-03 : 4. E-05 : 1.5E-02D: 1.1E-11
 3 : 3.8000001798719E+03 : 4.4E-06 : 6.3E-08 : 2. E-05C: 1.9E-11
 4 : 3.8000000000006E+03 : 1.5E-11 : 2.4E-13 : 3.6E-10D: 5.3E-10
```

The Newton iteration is initialized at $f_0(e^{i\theta}) = (29.6 + .1e^{i\theta}, -30.4 - .0001e^{i\theta} + .001(e^{i\theta})^2)$, which is near a local solution. Observe how the diagnostics OT1 and OT2 tend to zero at essentially quadratic rate. The same holds for the true error $|| f^k - f^* ||_{L^\infty_N}$.

Table 2

```
It :        Value        :||fk-f*||: Optimality Tests  :  Error
   :      SupGamma       :         :   OT1       OT2   :   NED
-----------------------------------------------------------------
 0 : 2.0428618081601E+04 : 1.4E+00 : 3.2E-02 : 2.9E-01 :   N/A
 1 : 2.0057600281609E+04 : 1.5E+00 : 5.7E-03 : 9.3E-02 : 3.1E-12
 2 : 2.0097292539204E+04 : 7. E-01 : 1.2E-02 : 3.6E-02D: 4.2E-11
 3 : 2.000858365994 E+04 : 3.1E-02 : 6.8E-04 : 3.4E-03C: 3.7E-11
 4 : 2.0000014412272E+04 : 4.8E-05 : 1.4E-06 : 6.6E-06D: 3.2E-11
 5 : 2.0000000000109E+04 : 7.6E-10 : 9.7E-12 : 7.3E-11C: 6.7E-11
```

In this example the Newton iteration converges quadratically to the constant function $f^*(e^{i\theta}) = (0, 0)$, which is not a local minimizer. The iteration was initialized with $f(e^{i\theta}) = (0.2 + 0.1e^{i\theta}, 1.4 + .1e^{i\theta} + 0.01e^{i\theta^2})$.

4.3 Nehari's problem

Let k be a function in L_N^∞. The distance of k to H_N^∞ is

$$\operatorname{dist}(k, H_N^\infty) := \inf_{f \in H_N^\infty} ||k - f||_{L_N^\infty} \tag{55}$$

and a *best* H^∞ *approximation to* k is a function $f_0 \in H_N^\infty$ that realizes the supremum in (55). Standard theory shows that f_0 exists. Moreover, if $f_0 \in \mathcal{C} + H_N^\infty$ then it is unique. We refer to the problem of calculating $\operatorname{dist}(k, H_N^\infty)$ and a minimizer $f_0 \in H_N^\infty$ as *Nehari's problem.*

Let $\mathcal{H}_k : H_1^2 \rightarrow H_N^2$ be the Hankel operator with action $\alpha \rightarrow P_{H_N^{2\perp}}[k\alpha]$. A theorem of Nehari for $N = 1$ asserts that the norm of the operator $\mathcal{H}_k$ satisfies

$$||\, \mathcal{H}_k \,||_{\mathcal{L}} \;=\; \operatorname{dist}(k, H_N^\infty) \tag{56}$$

while for $N > 1$ this follows from commutant lifting arguments. Another result is: if α_0 is a function at which $\mathcal{H}_k$ attains its norm and if f_0 is the best H^∞ approximation to k, then

$$k - f_0 = \frac{1}{\alpha_0}\mathcal{H}_k[\alpha_0] \tag{57}$$

Let $\mathcal{H}_k^*$ be the adjoint of $\mathcal{H}_k$, and set $\tau = \operatorname{dist}(k, H_N^\infty)$; then one can show that

$$\mathcal{H}_k^*[\, \mathcal{H}_k[\alpha_0]\,] = \tau^2 \alpha_0 \tag{58}$$

Equation (58) says that α_0 is a function associated with τ, the largest singular value of the compact operator $\mathcal{H}_k$ (i.e., the norm of $\mathcal{H}_k$). We will refer to equation (58) as the *Nehari - Commutant Lifting formula.*

4.4 Derivation of the Nehari-Commutant Lifting and other formulas

In this Subsection we derive relations (56), (57), and (58) directly from the optimality conditions in Theorem 2.

Note first that the Nehari problem is of the type OPT^∞, where

$$\Gamma(e^{i\theta}, z) = ||\; k(e^{i\theta}) - z\; ||_{C^N}^2 \tag{59}$$

To avoid technical issues we assume that $k \notin H_N^\infty$ and that $k \in H_N^\infty + \mathcal{C}$, hence there is a unique best H^∞ approximation f_0 to k. We also assume that $k - f_0$ is bounded away from 0. Thus Γ and f_0 satisfy the hypotheses of Theorem 2.

Set $\tau = \operatorname{dist}(k, H_N^\infty)$. The optimality conditions Ia. and Ib. of Theorem 2 in this case say that there exists λ measurable and positive on T, and a function $F \in H_N^1$ such that

$$\begin{aligned} ||k(e^{i\theta}) - f_0(e^{i\theta})||_{C^N}^2 &= \quad \tau^2 && (60)\\ \frac{1}{\lambda}(\overline{k - f_0}) &= \quad \chi F && (61) \end{aligned}$$

Let $\alpha_0 \in H^2_1$ be the outer spectral factor of $1/\lambda$. From (61) we have that

$$k - f_0 = \frac{1}{\alpha_0} \frac{\overline{\chi F}}{\overline{\alpha_0}} = \frac{1}{\alpha_0} P_{H_N^{2\perp}} [\frac{\overline{\chi F}}{\overline{\alpha_0}}] \tag{62}$$

Combining (61) and (62) one obtains

$$k - f_0 = \frac{1}{\alpha_0} P_{H_N^{2\perp}} [(k - f_0)\alpha_0] = \frac{1}{\alpha_0} P_{H_N^{2\perp}} [k\alpha_0] \tag{63}$$

This proves the formula (57).

We now check the identity (56). We trivially have that

$$||\mathcal{H}_k||^2 = ||\mathcal{H}_{k-f_0}||^2 \ \leq \ ||k - f_0||^2_{L^\infty_N} = \tau^2 \tag{64}$$

To prove the reverse inequality, note that by (60) the following relation holds for almost all θ:

$$\tau^2 = \frac{|\alpha_0(e^{i\theta})|^2 ||k(e^{i\theta}) - f_0(e^{i\theta})||^2_{C^N}}{|\alpha_0(e^{i\theta})|^2} = \frac{||\alpha_0(e^{i\theta})(k(e^{i\theta}) - f_0(e^{i\theta}))||^2_{C^N}}{|\alpha_0(e^{i\theta})|^2} \tag{65}$$

Combine (57) and (65) to obtain

$$\tau^2 = \frac{||\alpha_0(k - f_0)||^2_{L^2_N}}{||\alpha_0||^2_{L^2_N}} = \frac{||P_{H_N^{2\perp}}[\alpha_0(k - f_0)]||^2_{L^2_N}}{||\alpha_0||^2_{L^2_N}} \leq \ ||\mathcal{H}_k||^2_{\mathcal{L}} \tag{66}$$

and this finishes the proof of the equality (56).

At this point we are in position of deriving the Nehari - Commutant Lifting formula (58). Rewrite (61) as

$$\overline{\alpha_0}(\overline{k - f_0}) = (\chi/\alpha_0) \ F \tag{67}$$

and notice that by multiplying both members of (67) by $(k - f_0)^t$ and using (60) one obtains

$$\tau^2 \overline{\alpha_0} = (\chi/\alpha_0) \ (k - f_0)^t F \tag{68}$$

Now we reach a crossroads. The strategy behind our approach in Section 3 was to project (67) onto $H^{2\perp}_{N,0}$. In this case it produces a Toeplitz operator type formula, $P_{H^{2\perp}_{N,0}}[\ \overline{\alpha_0}\overline{k} - \overline{\alpha_0}\overline{f}\ ] = 0$, i.e., $\mathcal{T}_{k-f}[\ \alpha\ ] = 0$.

The strategy in Nehari and Commutant Lifting approaches is to project both terms in (67) onto $H^2_{N,0}$. This yields

$$P_{H^2_{N,0}}[\ \overline{\alpha_0}\overline{k}\ ] = (\chi/\alpha_0) \ F \tag{69}$$

Multiplying both members of (69) by $(k - f)^t$, using (68), and taking conjugate of both members we have

$$(\overline{k - f})^t P_{H_N^{2\perp}}[\ \alpha_0 k\ ] = \tau^2 \alpha_0 \tag{70}$$

Taking projection of both members of (70) onto H^2_N gives

$$P_{H^2_N}[\ \overline{k}^t P_{H_N^{2\perp}}[\ \alpha_0 k\ ]\ ] = \tau^2 \alpha_0 \tag{71}$$

Now equation (71) is precisely the Nehari - Commutant Lifting formula (58).

We see that the Nehari- Commutant Lifting formula is remarkable in that it puts together in a single equality both relations (60) and (61), while at the same time eliminates f and F (our approach in Section 3 eliminates τ and F). It is also notable that all this is accomplished with a formula that is *linear* in α. It is unfortunate that these properties do not hold for functions Γ more general than (59).

4.5 The Power Method for solving Nehari's Problem

A common way to solve Nehari's problem is to solve (58) for α_0 as an eigenvalue problem, and then solve for f_0 in (57). The power method is a natural choice since we are seeking the largest eigenvalue of the operator $\mathcal{H}_k^* \mathcal{H}_k$, and the corresponding eigenvector. It is a well known fact that, provided the multiplicity of this eigenvalue is equal to 1, the power method converges linearly to the solution.

The Power Method. Given k and a guess α to a solution to (58), update α to α_1 with

P1. $\alpha_+ := P_{H^2_N}[\, \overline{k}^t P_{H^{2\perp}_N}[\, k\alpha \,]\,]$.

P2. $\alpha_1 := \frac{1}{\|\ \alpha_+ \ \|_{L^2_1}} \alpha_+$.

Suppose that the unit circle is discretized to n (a power of 2) equally spaced points, so that one can calculate projections in $2n \log_2(n)$ operations by using a Fast Fourier Transform algorithm (we do not count additions). Then step P1 takes $N(4n \log_2(n) + 2n)$ operations, and step P2 takes Nn multiplications, 1 square root and Nn divisions. Thus the power method for solving Nehari's problem is of order $Nn \log_2(n)$. Also, note that memory requirements are low since no matrices have to be stored at any moment.

Newton iteration requires the inversion of a matrix M that arises from a discretization of the operator $T\prime_{f,beta}$ introduced in Section 3. The matrix M has a number of entries that scales with Nn^2, but it has special structure that can be exploited to reduce the operation count in the inversion (with algorithms of the type found in [SLK]), possibly reducing the order of the inversion to Nn^2. Also there exist algorithms for inverting a matrix by iteration which can be effective in this setting.

The power method is clearly efficient in many respects, and it is superior to our Newton iteration in memory requirements and operation count. Obviously it is also better when it comes to time actually used in calculations to achieve a solution. However, there is a crucial limitation of the power method: it is directly applicable only to very simple problems OPT^∞, namely Nehari problems.

The good qualities of the power method make it an interesting tool for solving OPT^∞problems through approximating Γ by Taylor expansions. This Taylor expansion is used to obtain a Nehari problem, or at least a problem close to it. This has been attempted before in [HMer1]. In Subsection 4.6 below the method of *Disk iteration*, introduced in [HMer1], is described.

4.6 Disk iteration for OPT^∞

We present here runs with the method described in [HMer1] for comparison with the runs output presented in Subsection 4.2. We refer to this method as *Disk iteration*, and it consists in the following step to update a current guess $f \in A_N$ at the solution.

Disk Iteration.

B1. Find $h = (h_1, h_2, \ldots, h_N) \in A_N$ that minimizes

$$\sup_{e^{i\theta}} \left\{ \Gamma(\cdot, f) + 2\Re\left(\frac{\partial \Gamma}{\partial z}(\cdot, f)^t h\right) + \sum_{k=1}^{N} \frac{\partial^2 \Gamma}{\partial z_k \partial \overline{z_k}}(\cdot, f)\ |h_k|^2 \right\}$$

B2. Perform a linesearch in $t > 0$ to find the best t_1 in $f + th$ according to a preset criterion.

B3. Set $f_1 = f + t_1 h$

A first glance at the algorithms reveals that the Taylor expansion of Γ about f, the current guess at the solution, given in Step B1 is not necessarily of the form of a Nehari problem. However, it can be solved by iterating on Nehari problems [HMer1]. Thus an implementation of the Disk iteration algorithm that uses the power method requires an inner loop to find a solution to Subproblem B1. We note that a modification of the expression in B1 may eliminate the inner loop at the expense of *all* the second order information [HMer1].

Another feature of Subproblem B1 is that the Taylor expansion it contains does not capture all second order information from most functions Γ. One expects that this would imply a linear order of convergence for the algorithm. This is indeed the case in the vector valued case $N > 1$, but it is really surprising that (in practice) quadratic order of convergence occur in the scalar case $N = 1$ (see [HMer1]).

The criterion in Step B2 can take different forms. A natural one is to minimize $\sup_\theta \Gamma(e^{i\theta}, f(e^{i\theta}) + th(e^{i\theta}))$. We shall see (in Table 3 below) that this choice may lead to unsatisfactory progress in diagnostics such as OT2. One can attempt to remedy this by performing the linesearch to minimize $OT_2(f + th)$ in t if needed. It is known [HMer1] that the Disk iteration gives quadratic convervence rate in practice when $N = 1$, and no better than linear rate when $N > 1$.

Tables 3, 4 and 5 illustrate the behavior of the Disk iteration when applied to the function Γ in (54).

Table 3

```
It :        Value         :||fk-f*||: Optimality Tests  :  Error
   :      SupGamma        :         :   OT1       OT2   :   NED
-----------------------------------------------------------------
 0 : 3.8165268771511E+03 : 5.7E-01 : 8.3E-03 : 1.6E+00 :   N/A
 1 : 3.8042136191184E+03 : 1.8E-01 : 1.7E-03 : 1.7E+00D: 3.1E+00
 2 : 3.8022080149394E+03 : 1.6E-01 : 7.6E-04 : 1.8E+00C: 8.9E-03
 3 : 3.8016053529266E+03 : 1.3E-01 : 5.5E-04 : 1.4E+00 : 1. E-03
 4 : 3.8012267267355E+03 : 1.2E-01 : 4.2E-04 : 1.1E+00 : 6.1E-04
 5 : 3.8008576491564E+03 : 10.E-02 : 2.9E-04 : 1. E+00 : 4.8E-04
 6 : 3.8006645200494E+03 : 9. E-02 : 2.2E-04 : 9.4E-01D: 4.3E-04
 7 : 3.8004623215894E+03 : 7.8E-02 : 1.5E-04 : 7.8E-01C: 3.9E-04
 8 : 3.8003710328874E+03 : 7. E-02 : 1.2E-04 : 7. E-01D: 3.3E-04
 9 : 3.8002696884688E+03 : 6.2E-02 : 7.9E-05 : 5.9E-01C: 3.1E-04
10 : 3.8002140361056E+03 : 5.5E-02 : 6.5E-05 : 5.3E-01D: 2.6E-04
11 : 3.8001623242094E+03 : 4.9E-02 : 4.3E-05 : 4.6E-01C: 2.5E-04
12 : 3.8001285408615E+03 : 4.4E-02 : 3.8E-05 : 4.1E-01D: 2.1E-04
13 : 3.8001008070299E+03 : 3.9E-02 : 2.4E-05 : 3.7E-01C: 2. E-04
14 : 3.8000810369876E+03 : 3.5E-02 : 2.3E-05 : 3.3E-01D: 1.6E-04
15 : 3.8000684895455E+03 : 3.2E-02 : 1.5E-05 : 3. E-01C: 1.6E-04
16 : 3.8000540353996E+03 : 2.8E-02 : 1.7E-05 : 2.6E-01D: 1.3E-04
```

The Disk iteration with linesearch on the value of Γ and initialized at $f_0(e^{i\theta}) = (29.6 + .1e^{i\theta}, -30.4 - .0001e^{i\theta} + .001(e^{i\theta})^2)$, gives very slow progress towards the solution. Observe that the diagnostic OT2 decreases at much slower rate than OT1. This is common behavior for this algorithm when optimizing over A_2. Also note that the numerical error diagnostic remains relatively large during the iteration. Compare with Table 1, where the function Γ and the initialization in f are the same as those in the present run, and quadratic convergence rate is achieved.

Table 4

```
It :        Value        :  Step   : Optimality Tests  :  Error  :
   :      SupGamma       :         :   OT1       OT2   :   NED   :
-------------------------------------------------------------------
 0 : 3.8165268771511E+03 :   N/A   : 8.3E-03 : 1.6E+00 :   N/A   :
 1 : 3.8033527812023E+03 : 6.8E-01 : 1.4E-03 : 1.6E+00D: 3.1E+00 :
 2 : 3.8020614622957E+03 : 3.2E-01 : 7.2E-04 : 1.8E+00C: 1.1E-04 :
 3 : 3.8015296776761E+03 : 2.8E-01 : 5.2E-04 : 1.1E+00D: 3.2E-06 :
 4 : 3.8010704285019E+03 : 2.4E-01 : 3.6E-04 : 1.1E+00C: 5.9E-07 :
 5 : 3.8007798726506E+03 : 2.2E-01 : 2.5E-04 : 1. E+00 : 1.8E-06 :
 6 : 3.8005523537358E+03 : 1.9E-01 : 1.8E-04 : 9.2E-01D: 4.4E-07 :
 7 : 3.8004021530084E+03 : 1.7E-01 : 1.2E-04 : 7.7E-01C: 1.2E-06 :
 8 : 3.8002935927643E+03 : 1.5E-01 : 8.5E-05 : 6.8E-01D: 3.4E-07 :
 9 : 3.8002183486351E+03 : 1.3E-01 : 6. E-05 : 5.8E-01C: 5.3E-07 :
10 : 3.8001658443452E+03 : 1.2E-01 : 4.2E-05 : 5.2E-01D: 2.2E-07 :
11 : 3.8001269523431E+03 : 1. E-01 : 3. E-05 : 4.6E-01C: 1.3E-06 :
12 : 3.8000993556187E+03 : 9.3E-02 : 2.1E-05 : 4. E-01D: 2. E-07 :
13 : 3.8000780732537E+03 : 8.3E-02 : 1.5E-05 : 3.6E-01C: 4.4E-07 :
14 : 3.8000625135672E+03 : 7.4E-02 : 1.1E-05 : 3.2E-01D: 1.8E-07 :
15 : 3.8000501876633E+03 : 6.7E-02 : 8. E-06 : 2.9E-01C: 6.4E-07 :
16 : 3.8000408181867E+03 : 6. E-02 : 5.9E-06 : 2.6E-01D: 1.4E-06 :
```

The example shown in Table 3 is now run with automatic smoothing on the current guess at the solution. This lowers the numerical error NED as the iteration progresses and improves OT1, but OT2 is practically unchanged. Thus numerical noise is not the source of the lack of progress in the run in Table 1.

Table 5

```
It :         Value       :||fk-f*||: Optimality Tests  :  Error
   :       SupGamma      :         :   OT1       OT2   :   NED
-----------------------------------------------------------------
 0 : 3.8165268771511E+03 : 5.7E-01 : 8.3E-03 : 1.6E+00 :   N/A
 1 : 3.8014192462609E+03 : 1.4E-01 : 5.6E-04 : 1.1E+00D: 3.1E+00
 2 : 3.8004689130398E+03 : 1.3E-02 : 3. E-04 : 5.4E-02D: 1.1E-02
 3 : 3.8001160653124E+03 : 4.6E-03 : 6.2E-05 : 4. E-02C: 1.1E-03
 4 : 3.8000022236943E+03 : 3. E-03 : 9.8E-07 : 3.3E-02D: 8.4E-05
 5 : 3.8000007986885E+03 : 1.1E-03 : 4.5E-07 : 1.1E-02C: 3.3E-06
 6 : 3.8000001592016E+03 : 7.1E-04 : 1.2E-07 : 7.6E-03D: 1.5E-06
 7 : 3.8000000629256E+03 : 1.3E-04 : 3.7E-08 : 1.4E-03D: 8.7E-07
 8 : 3.8000000028458E+03 : 1.1E-04 : 1.9E-09 : 1.2E-03C: 4.9E-05
 9 : 3.8000000007779E+03 : 5.9E-06 : 5.3E-10 : 6.8E-05D: 4.3E-05
10 : 3.8000000002904E+03 : 3.5E-06 : 2.1E-10 : 4.1E-05C: 2.6E-06
11 : 3.8000000000528E+03 : 4.1E-07 : 4.1E-11 : 4.9E-06D: 1.6E-06
12 : 3.8000000000284E+03 : 1.9E-07 : 1.9E-11 : 2.3E-06C: 2. E-07
13 : 3.8000000000045E+03 : 3.1E-08 : 3.5E-12 : 3.8E-07D: 9.8E-08
14 : 3.8000000000023E+03 : 4.2E-09 : 1.2E-12 : 5.3E-08C: 1.8E-08
```

The example shown in Table 3 is now run with a mixed strategy in the linesearch: if f is the current guess at the solution and h is the update direction generated by the algorithm, then either $OT_2(f+th)$ or the supremum of $\Gamma(\cdot, f+th)$ is minimized in t, depending on progress with respect to previous iterations. The progress improves to a clear linear rate of convergence. Compare with Table 1.

5 Algorithms for solving OPT^1

In this section we discuss algorithms for solving OPT^1. Newton algorithm is considered first. We show that it is locally quadratically convergent in Sobolev space norm $||\cdot||_{\mathcal{S}_p}$ to strict local directional solutions f^* to OPT^1 when f^* is smooth. The convergence is illustrated with output of a computer run.

We then consider a class of descent algorithms based upon Taylor expansions of the function Γ. No convergence result is given for these algorithms. Rate of convergence of these algorithms is expected to be inferior to Newton iteration. The justifications for considering the descent algorithms are (1) the requirements of computer memory are low, and (2) no inversion of operator is required. This makes the descent algorithms a practical alternative to Newton iteration in many large problems.

5.1 The Newton iteration for solving OPT^1

We assume that we are given a smooth function Γ and we wish to find a minimizer f^* for OPT^1. We take advantage of the hindsight gained in Section 3 and skip some detail in the calculations.

Let Q be the operator defined on $\mathcal{A}_N^\infty$ by

$$Q[f] = P_{H^{2\perp}_{N,0}}[\frac{\partial\Gamma}{\partial z}(\cdot,f)] \tag{72}$$

By Theorem 1, the function f^* for has to satisfy the equation

$$Q[f^*] = 0 \tag{73}$$

We now attempt to solve equation (73) iteratively. To update a guess f to $f+h$ using Newton 's iteration the function h must be a solution to the equation

$$Q[f] + Q'_f[h] = 0 \tag{74}$$

Thus if the linear operator Q'_f is invertible, the function h satisfies $h = -(Q'_f)^{-1}[Q[f]]$. Note that if

$$A = \frac{\partial^2\Gamma}{\partial z\partial\overline{z}}(\cdot,f) \quad \textit{and} \quad B = \frac{\partial^2\Gamma}{\partial^2 z}(\cdot,f) \tag{75}$$

then an explicit formula for the operator Q'_f is

$$Q'[h] = P_{H^{2\perp}_{N,0}}[A\overline{h} + Bh\ ] = (\mathcal{C}\mathcal{T}_A + \chi\mathcal{H}_{\overline{\chi}B})\,[\ h\ ] \tag{76}$$

Proposition 7 *Let Γ be smooth and strictly plurisubharmonic in z, and let $f^* \in \mathcal{A}_N^\infty$ be a strict local directional solution to OPT^1. Then there exists a neighborhood $V \subset H^2_N$ of f^* such that for every $f \in V \cap \mathcal{A}_N^\infty$ the operator Q'_f is bounded and invertible from $\mathcal{S}_p \cap H^2_N$ to $\mathcal{S}_p \cap H^{2\perp}_{N,0}$. Moreover, if h is in $\mathcal{A}_N^\infty \cap H^{2\perp}_{N,0}$ then $(Q'_f)^{-1}[\ h\ ]$ is in $\mathcal{A}_N^\infty$.*

Proof. Let f be any function in $\mathcal{A}_N^\infty$. Since Γ is strictly p.s.h. in z, the smooth function A in (75) is never 0, thus it is uniformly invertible. Then the operator $\mathcal{T}_A$ is invertible by Lemma 4. Since

$\mathcal{H}_B$ (with B given as in (75)) is compact, it follows from representation (76) that Q'_f is a Fredholm operator with index 0 that maps $\mathcal{S}_p \cap H^2_N$ into $\mathcal{S}_p \cap H^{2\perp}_N$.

To obtain the conclusion of the Theorem is enough to prove that Q'_{f^*} has trivial kernel when defined on the space H^2_N, since then one can proceed as in the proof to Theorem 6 to obtain the set W in the statement of the Theorem. In contrast to our results in Section 3, proving that Q'_{f^*} has trivial kernel turns out to be much simpler in the present setting.

If $Q'_{f^*}[\ h\ ] = 0$, then there exists a function $G \in H^2_N$ such that

$$A\overline{h} + Bh = \chi G \tag{77}$$

Therefore we have the relation

$$\int_0^{2\pi} h^t A\overline{h} + \Re h^t B h \frac{d\theta}{2\pi} = \int_0^{2\pi} \chi h^t G \frac{d\theta}{2\pi} \tag{78}$$

The right hand side member of (78) is 0. But since f^* is a strict local directional solution, the left hand side of (78) is 0 only when $h = 0$, by Theorem 1. Thus Q'_{f^*} has trivial kernel. ■

Corollary 2 *Fix $p \geq 0$. Then there exists a constant $c > 0$ and $V \subset \mathcal{A}^\infty_N$ a neighborhood of f^* in the norm of $\mathcal{S}_p$ such that whenever $f \in V$, the function $\tilde{f} := f - (Q'_f)^{-1}[\ Q[\ f\ ]]$ is a well defined element of A^∞_N and it satisfies*

$$\|\ \tilde{f} - f^*\ \|_{\mathcal{S}_p} \leq\ c\ \|\ f - f^*\ \|_{\mathcal{S}_p} \tag{79}$$

The computer output in Table 6 illustrates the behavior of the Newton iteration for solving OPT^1 when Γ is given by (54).

Table 6

```
It :        Value          :  Step   : Opt Test:  Error  :
   :      L1 Gamma         :         :   OT3   :   NED   :
---------------------------------------------------------
 0 : 3.7987529455866E+03 :   N/A   : 1.6E+00 :   N/A   :
 1 : 3.797381652889 E+03 : 4.4E-01 : 1. E-02 : 2.5E+01 :
 2 : 3.7973793411253E+03 : 4.6E-02 : 1.7E-04 : 4.2E-01 :
 3 : 3.7973793411034E+03 : 3.9E-05 : 1.7E-10 : 7.2E-04 :
```

Newton iteration for OPT^1 when Γ is given by (54). The initial guess is $f_0(e^{i\theta}) = (29.6 + .1e^{i\theta}, -30.4 - .0001e^{i\theta} + .001(e^{i\theta})^2)$. The grid on the unit circle for discretization has 128 points. The column OT3 corresponds to $OT_3(f) = ||P_{H^{2\perp}_N}[\frac{\partial\Gamma}{\partial z}(\cdot, f)]||_{L^2}$, where f is the current guess at the optimizer function.

5.2 Descent algorithms for solving OPT^1

Consider Γ as in Theorem 1, $f \in A_N$, and N by N- continuous matrix valued functions A_0 and B_0 on T (which may or may not depend on f), such that A_0 takes self-adjoint values, and B_0 takes symmetric values. Now define $\Gamma_{A_0,B_0,f}$ by

$$\Gamma_{A_0,B_0,f}(e^{i\theta},z) = \Gamma(e^{i\theta}, f(e^{i\theta})) + 2\Re\frac{\partial\Gamma}{\partial z}(e^{i\theta}, f(e^{i\theta}))^t z + \overline{z}^t A_0(e^{i\theta})z + \Re z^t B_0(e^{i\theta})z. \tag{80}$$

Observe that $\Gamma_{A_0,B_0,f}$ *coincides with the Taylor expansion of* Γ *about* f *up to (at least) order 1.* It represents the order 2 expansion if A_0 and B_0 are given by (18).

The algorithms are based on the following step for producing an update f_1 to a given $f \in A_N$:

STEP.

A1. Find $h = h_1 \in A_N$ that minimizes

$$\int_0^{2\pi} \Gamma_{A_0,B_0,f}(e^{i\theta}, h(e^{i\theta}))\frac{d\theta}{2\pi}$$

A2. Find t_1 that minimizes $||\Gamma(\cdot, f + th_1)||_{L^1}$

A3. Set $f_1 = f + t_1 h_1$

Proposition 8 *If the function* $det(A_0)$ *is uniformly bounded away from zero on* ∂D*, the equation (28) has only trivial solutions, and*

$$\int_0^{2\pi} \overline{h}^t A_0 h + \Re\, h^t B_0 h \frac{d\theta}{2\pi} > 0, \qquad \forall h \in H^2_N,\ h \neq 0, \tag{81}$$

then solutions h^* *to the subproblem* **A1** *exist in* H^2_N *and are unique.*

Proof: The optimality condition I. of Theorem 1 says that solutions to the subproblem must satisfy the equation

$$P_{H^{2\perp}_{N,0}}\left(a_0 + A_0\overline{\phi} + B_0\phi\right) = 0 \tag{82}$$

The hypotheses on A_0 and Lemma 4 imply that the operator $\phi \rightarrow P_{H^{2\perp}_{N,0}}\left(A_0\overline{\phi} + B_0\phi\right) = 0$ is a Fredholm operator in $\Phi(H^2_N, H^{2\perp}_N)$, that has index 0. It is also one-to-one by Lemma 6, hence it is invertible. Thus the equation (82) has a solution $h^* \in H^2_N$, and it is unique. By hypothesis, the condition II of Theorem 1 is satisfied with strict inequality by all nonzero $h \in H^2_N$. By Theorem 1, h^* is a solution to subproblem **A1**. ■

In many cases it is possible to write down h_1 in **A1** explicitly.

Proposition 9 *Let* Q *be the outer analytic spectral factor of* A_0*. If* $B_0(e^{i\theta}) = 0$*, for all* $e^{i\theta}$*, then the solution* $h_1 \in H^2_N$ *to Subproblem* **A1** *is given by*

$$h_1 = -Q^{-1}P_{H^2_N}\left(\frac{\overline{\partial\Gamma}}{\partial z}(\cdot, f)Q^{-1}\right) \tag{83}$$

Proof: Complete the squares in (80) and use the identity $A_0 = \overline{Q^t}Q$ to obtain the identity

$$\begin{aligned} \Gamma_{A_0,B_0,f}(e^{i\theta}, z) \quad &= \Gamma(e^{i\theta}, g(e^{i\theta})) - \left\| \overline{\frac{\partial\Gamma}{\partial z}}(e^{i\theta}, f(e^{i\theta}))Q^{-1}(e^{i\theta}) \right\|^2_{C^N} \\ &+ \left\| \frac{\overline{\partial\Gamma}}{\partial z}(e^{i\theta}, f(e^{i\theta}))Q^{-1}(e^{i\theta}) + Q(e^{i\theta})z \right\|^2_{C^N} \end{aligned} \tag{84}$$

Formula (83) follows immediately from (84). ∎

The class of algorithms defined with the Step above has as a special case an infinite dimensional (complex) version of the Gauss-Newton method for minimizing non-linear functionals in the L^2 sense (See [GMW], [W]). The problem is

Problem. Given a function $G : \partial D \times C^N \longrightarrow C^N$, find $f \in H^2_N$ that minimizes

$$\int_0^{2\pi} ||G(e^{i\theta}, f(e^{i\theta}))||^2_{C^N} \frac{d\theta}{2\pi} \tag{85}$$

In standard Gauss-Newton L^2 optimization one solves a subproblem of the form

$$\inf_h ||G(\cdot, f) + DG_f[h]||_{L^2} \tag{86}$$

where DG_f represents the Jacobian of G at f; thus

$$DG_f[z] = a_z^t z + a_{\overline{z}}^t \overline{z}$$

where $a_z = \frac{\partial G}{\partial z}(\cdot, f(\cdot))$ and $a_{\overline{z}} = \frac{\partial G}{\partial \overline{z}}(\cdot, f(\cdot))$. It is clear that elementary calculations with $||G(\cdot, f) + DG_f[h]||^2_{C^N}$ yield A_0 and B_0 in the representation (80) for the integrand in (85).

6 References

[A] T. Ando. On a pair of commutative contractions, Acta Sci. Math., 24 (1963), 88-90

[AAK1] V. M. Adamjan, D. Z. Arov, and M. G. Krein. Infinite Hankel matrices and generalized problems of Caratheodory-Fejer and F. Riesz. Functional Anal. Appl. 2(1968), 1-18.

[AAK1] V. M. Adamjan, D. Z. Arov, and M. G. Krein. Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. Math. USSR Sbornik, vol. 15, 1972, pp. 31-73.

[AAK2] V. M. Adamjan, D. Z. Arov, and M. G. Krein. Infinite block Hankel matrices and related extension problems. AMS Transl., vol. 111, 1978, pp. 133-156.

[BB] C. G. Barrat and S. P. Boyd, *Linear Controller Design*, Prentice-Hall, 1991.

[BHM] J. Bence, J. W. Helton, and D. Marshall. H^∞ optimization. Proc., Conference on Decision and Control, Athens, 1986.

[BS] A. Bottcher, and B. Silbermann. Analysis of Toeplitz Operators. Springer Verlag, Berlin, 1990.

[Dym] H. Dym, J Contractive Matrix Functions Reproducing Kernel Hilbert Spaces and Interpolation. Regional Conferences Series in Mathematics, **71**, 1989.

[F] Computing Optimal H^∞ -Controllers. In: H^∞ -Control Theory. Ed. E. Mosca, L. Pandolfi. Lecture Notes in Mathematics vol. 1496, Springer Verlag, Berlin 1991.

[FF] C. Foias and A. Frazho, *The commutant lifting approach to interpolation problems*, Operator Theory: Advances and Applications, vol. 0T44, 1990.

[Fr] B. A. Francis, *A Course in H^∞ Control Theory*, Lecture Notes in Control and Information Sci., vol 88, Springer-Verlag, 1986.

[GM] K. Glover and D. C. McFarlane, “Robust controller design using normalized coprime factor plant descriptions,” in *Lecture Notes in Control and Information Sciences*, **138**, Springer-Verlag, 1990.

[Hs] H. Helson. Lectures on Invariant Subspaces. Academic Press, New York, 1963.

[H] J. W. Helton, et al., Operator Theory, Analytic Functions, Matrices, and Electrical Engineering, in *Regional Conference Series in Mathematics*, no. **68**, American Mathematical Society, 1987.

[HM] J. W. Helton and D. Marshall, Frequency domain design and analytic selections. Indiana Journal of Math.

[HMer1] J. W. Helton and O. Merino. *A Novel Approach to Accelerating Newton's Method for Sup-Norm Optimization Arising in H^∞ Control.* To appear in Journal of Optimization Theory and Applications.

[HMer2] J. W. Helton and O. Merino, Conditions for Optimality over H^∞. To appear publication in SIAM Journal on Control and Optimization.

[HS] J. W. Helton, and D. Schwartz, Analytic best approximation to vector valued functions, preprint.

[Mer] On Qualitative Properties of Solutions to Problems of Optimization Over H^∞. To appear in American Journal of Mathematics.

[NF] B. Sz. Nagy, and C. Foias. Harmonic Analysis of Operators on Hilbert space. North Holland, New York, 1970.

[RR1] M. Rosenblum, and J. Rovnyak, *Hardy Classes and Operator Theory*, Oxford University Press, 1985

[RR2] M. Rosenblum, and J. Rovnyak, The factorization problem for nonnegative operator valued functions, Bull., Amer. Math. Soc., **77** (1971), 287–318.

[S] D. Sarason. Generalized interpolation in H^∞. Trans. Amer. Math. Soc., 127 (1967), 179-203.

[SLK] A.H. Sayed, H. Lev-Ari, and T. Kailath Fast Triangular Factorization of the Sum of Quasi-Toeplitz and Quasi-Hankel Matrices, Oct 21, 1992, To appear Linear Algebra and Its applications.

[W] G. A. Watson. Approximation Theory and Numerical Methods. J. Wiley, New York, 1980.

[Yng] N. Young, *An Introduction to Hilbert Space*, Cambridge University Press, 1988.